WPS4372

POLICY RESEARCH WORKING PAPER 4372

Public Expenditure and Growth

Santiago Herrera

The World Bank
Development Economics Vice-Presidency
October 2007

Policy Research Working Paper 4372

Abstract

Given that public spending will have a positive impact on GDP if the benefits exceed the marginal cost of public funds, the present paper deals with measuring costs and benefits of public spending. The paper discusses one cost seldom considered in the literature and in policy debates, namely, the volatility derived from additional public spending. The paper identifies a relationship between public spending volatility and consumption volatility, which implies a direct welfare loss to society. This loss is substantial in developing countries, estimated at 8 percent of consumption. If welfare losses due to volatility are this sizeable, then measuring the benefits of public spending is critical.

Gauging benefits based on macro aggregate data requires three caveats: a) considering of the impact of the funding (taxation) required for the additional public spending; b) differentiating between investment and capital formation; c) allowing for heterogeneous response of output to different types of capital and differences in network development. It is essential to go beyond country-specificity to project-level evaluation of the benefits and costs of public projects. From the micro viewpoint, the rate of return of a project must exceed the marginal cost of public funds, determined by tax levels and structure. Credible evaluations require microeconomic evidence and careful specification of counterfactuals. On this, the impact evaluation literature and methods play a critical role. From individual project evaluation, the analyst must contemplate the general equilibrium impacts. In general, the paper advocates for project evaluation as a central piece of any development platform. By increasing the efficiency of public spending, the government can permanently increase the rate of productivity growth and, hence, affect the growth rate of GDP.

This paper—a product of theDevelopment Economics Vice-presidency—is part of a larger effort in the department to analyze the impact of public expenditure on growth and other outcomes, such as human and physical capital formation. Policy Research Working Papers are also posted on the Web at http://econ.worldbank.org. The author may be contacted at sherrera@worldbank.org.

Public Expenditure and Growth

Santiago Herrera[1]

[1] Lead Economist in the Development Economics Vice-presidency (DEC), The World Bank. A version of this paper was presented in the G-20 Workshop on Fiscal Policy, in Istanbul, Turkey. The author thanks Sara Calvo, Sebastien Dessus, Edgardo Favaro, Jorge Garcia, Gerhard Glomm, Jeff Lewis, Eduardo Ley, Ramon Moreno, Claudia Sepulveda, Luis Serven, Kostas Tsatsaronis, and Jan von der Goltz for comments to a previous draft. The findings, interpretations and conclusions are the author's responsibility and do not necessarily represent the view of the World Bank or its Executive Directors.

I. Introduction

Public spending influences output growth if it affects capital formation or productivity growth. That impact may be direct or indirect. The direct one is mediated by the effectiveness and efficiency with which spending is transformed into capital. The policymaker and other stakeholders need to make operational the "efficiency" and "effectiveness" notions for different reasons. The policymaker has to choose among alternative projects and needs to gauge the impact of public spending on outcomes to make changes if the objectives are not being met. Other stakeholders will evaluate efficiency as part of their assessment of the delegation contract subscribed with their representatives to guide spending and taxation.

The indirect impact of public spending on growth operates through people's production, consumption and labor supply reactions to changes in relative prices and income derived from public spending variations. Among these effects, those derived from the taxation needed to fund additional spending are of particular relevance, given the inseparability of spending decisions and taxation. From a macro standpoint, the inseparability arises from the requirement that fiscal policy be sustainable. From the micro perspective, spending and taxation are not independent because the cost-benefit analysis of any public project is, essentially, a comparison of its expected social benefit with the marginal social cost of public funds, which is society's cost of raising an additional dollar of public revenue.

Besides its impact on capital formation, public spending may also affect productivity growth through changes in the efficiency of resource use. In this respect, Arnold Harberger (2004) recently described the revolution he had witnessed in policymaking during the last four decades. He noted how, on trade policy, the world had moved from a high-tariff protectionist environment toward a more liberalized trading system. On exchange rate policy, the free floating exchange rate and the monetary union paradigms had eased the action of market forces in determining the real value of currencies. In monetary policy, independent central banks or monetary unions had significantly altered the ways in which countries managed their currencies. On public expenditure management, however, he noted how most countries continued practicing it as several decades ago, with little regard for economic efficiency. Among several reasons, he pointed to the lack of technical analysis in the measurement of costs and benefits in the practice of expenditure management.

This paper examines the impact of public spending on growth and welfare, and discusses its effectiveness in achieving outcomes. The analysis can be done from a macro perspective or a micro standpoint. While the macro approach concentrates on aggregate public spending and its welfare effects on society as a whole, the micro one focuses on the impact of specific programs on individuals or groups of them. The methods and data sources for each are different. But both coincide in the need of clearly delineating the objectives of public spending and the requirement of measuring outcomes to assess whether the objectives were met or not. In general, the paper endorses the proposal of expenditure evaluation and monitoring as a development platform.

Along these lines, the paper's first two sections discuss macro and micro aspects of the impact of public spending on growth and welfare. The first section, on macroeconomics, explores the dual possibility of public spending affecting output via its impact on capital formation or other production inputs, or affecting directly society's welfare. The first option, the "productive" aspect of spending, is the most explored in the literature that focuses on quantifying the productivity of public spending. The section overviews the literature and highlights several challenges that policymakers face when drawing policy lessons from it.

The second option, of public spending directly affecting welfare may come through different channels. The one generally explored in the literature derives from the provision of goods or services that are strictly utility-enhancing. The paper abstracts from this channel, and focuses on people's risk aversion and the impact that public spending has on the volatility of their consumption. Given the pervasiveness of volatility in developing countries and their vulnerability to it, the section highlights the association between public spending and the volatility of consumption. Given people's preference for smooth consumption, there is a substantial welfare loss in developing countries, setting an additional hurdle for any expected benefits derived from public spending.

The second section, on microeconomics, focuses on the evaluation of public projects and the impact assessment of specific programs on capital formation. Here too, spending and taxation are inseparable because economic efficiency requires that a project's expected social benefit exceeds the marginal cost of public funds. Although in theory the economic efficiency criterion seems simple, in practice there are major challenges, especially quantifying the benefits of a program. On this, the recent impact evaluation literature and methods can illuminate the way forward. For policy purposes, it is essential to be able to extrapolate results from individual program evaluations to other settings or to different scales of operation of it, highlighting the need of considering its general equilibrium effects. This brings back the analysis to macro modeling and completes the evaluation cycle of the effect of public spending.

II. Macro Aspects of Public Spending

This section focuses on the long run growth effect of public spending, and highlights caveats to bear in mind when drawing policy lessons from macro econometric work that quantifies the productivity of public spending. The particular long-run aspects highlighted in this section derive from the effects of public spending volatility on capital formation and household consumption. The first effect on investment has been explored, but the effect on consumption volatility has not. Given individual's preferences for smooth consumption, this entails a direct welfare loss to society, quantified at substantial levels in developing economies.

The section simplifies the analysis in many ways. For instance, it abstracts from the short run demand-side effects of public spending on economic activity; it also condenses critical discussions of the composition of public spending by considering a simple parameter, the elasticity of output with respect to public spending or public capital. Much

of the section is an overview of the challenges of measuring the productivity of spending and its impact on capital formation based on aggregate macro data.

A. Expenditure volatility impact on capital formation and welfare

When public spending grows beyond the long-run sustainable funding level provided by taxation, it is subject to the volatility of financial markets. Though volatility may be cushioned by public debt management in the short run, in the long run disequilibria between taxation and expenditure lead to policy instability, as revenue or spending have to be adjusted. This mismatch produces stop-go cycles of public spending, generating fiscal policy volatility that compounds economic uncertainty arising from other sources. Not surprisingly, Table 1 shows how public spending is more volatile in poor and heavily indebted countries that rely on external saving to close financial gaps. Another interesting fact depicted in Table 1 is that volatility decreases as the income level rises, reaching a minimum in the subset of the European Monetary Union countries within the high income group.

Not only is public spending volatility higher in absolute terms in less developed nations, but it is also larger relative to that of household consumption or GDP, as Table 1 shows. This fact by itself would not be problematic if it reflected active use of a policy instrument with the target variable (output or consumption) being stabilized. In these circumstances, fiscal policy, and in particular public spending, could play a positive role within the tools disposable to the policymaker to stabilize output. But the procyclical nature of fiscal policy in most developing countries, explored elsewhere (Alesina and Tabellini,2005; Talvi and Vegh, 2005), suggests that fiscal policy may be destabilizing. More importantly, and in the same direction, ther is a positive association between public spending volatility and output and consumption volatility as shown in Figures 1 and 2.[2]

[2] Volatility is defined as the standard deviation of the per-capita GDP and consumption per capita growth rates during 1961-2005.

Table 1
Government Expenditure* Growth and Volatility across the World, by level of income 1961- 2005
(growth rates)

	Expenditure growth			**Volatility** relative to that of :	
	(1)	(2)	(3)=(2)/(1)		
	$\bar{x}$	σ	$\sigma / \bar{x}$	Household consumption	GDP
Heavily indebted poor countries (HIPC)	1.42	5.06	3.56	2.58	2.80
High income	2.78	1.29	0.46	0.91	0.83
European Monetary Union	2.87	1.25	0.43	0.66	0.70
Least developed countries: UN classification	2.00	5.51	2.76	2.00	2.82
IDA only	1.97	4.14	2.10	2.24	1.80
Low income	5.25	4.52	0.86	2.27	2.25
Middle income	4.60	2.91	0.63	1.58	1.72
World	**3.00**	**1.23**	**0.42**	**0.98**	**0.89**

*Refers to general government consumption expenditure
** Standard deviation of government expenditure growth relative to the standard deviation in the growth of each variable.
Source: Author's calculations based on WDI.

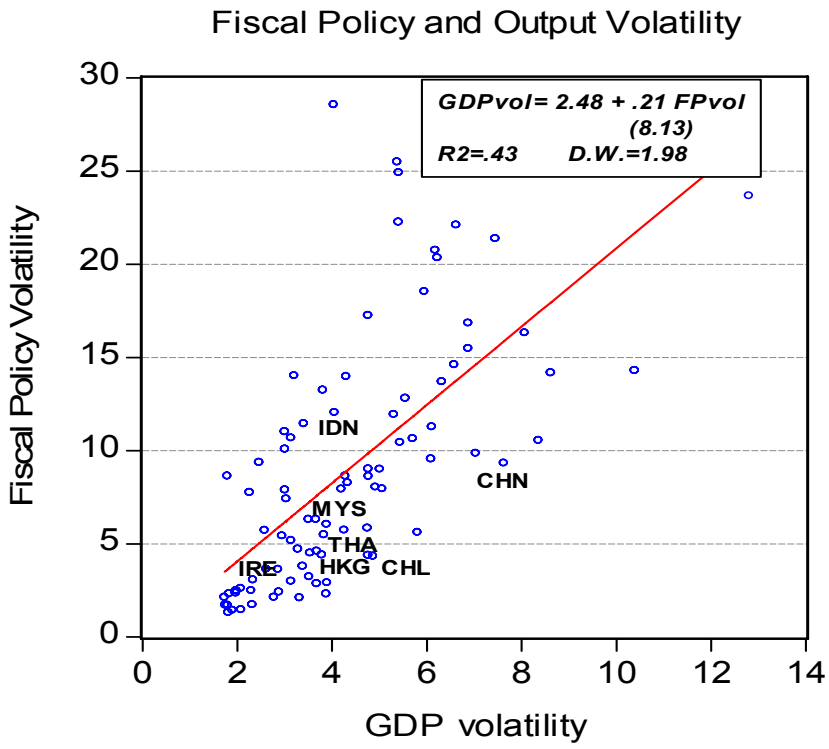
Output and consumption volatilities show different patterns in high income and low income countries. The volatility of output and consumption is greater in lower income countries, and, consumption is more volatile than income in all countries except in the high income category (Table 1A). Household consumption is extremely volatile in developing countries and that implies a substantial welfare cost for society.

Table 1A
Output and Consumption across the World, by level of income 1961- 2005
(per capita growth rates)

	GDP			**Consumption**			
	(1)	(2)	(3)=(2)/(1)	(4)	(5)	(6)=(5)/(4)	(7)=(5)/(2)
	$\bar{x}$	σ	$\sigma / \bar{x}$	$\bar{x}$	σ	$\sigma / \bar{x}$	σ con/σ gdp
Heavily indebted poor countries (HIPC)	0.13	1.80	13.57	-0.25	2.06	-8.09	1.15
High income	2.61	1.58	0.61	2.54	1.36	0.54	0.86
Least developed countries: UN classification	0.44	1.93	4.43	-0.46	2.70	-5.90	1.40
Low income	1.95	1.99	1.02	1.12	1.98	1.77	1.00
Middle income	2.73	1.67	0.61	2.70	1.72	0.64	1.03
World	**1.92**	**1.38**	**0.72**	**1.81**	**1.12**	**0.62**	**0.81**

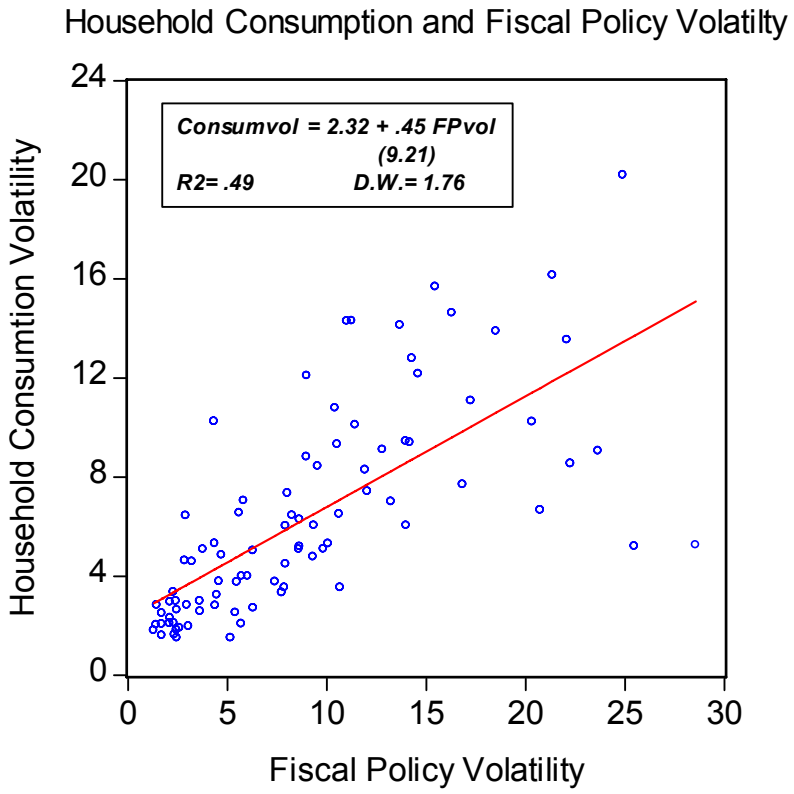
Source: Author's calculations based on WDI

Figure 1



Source: Author's calculations

Figure 2



Source: Author's calculations

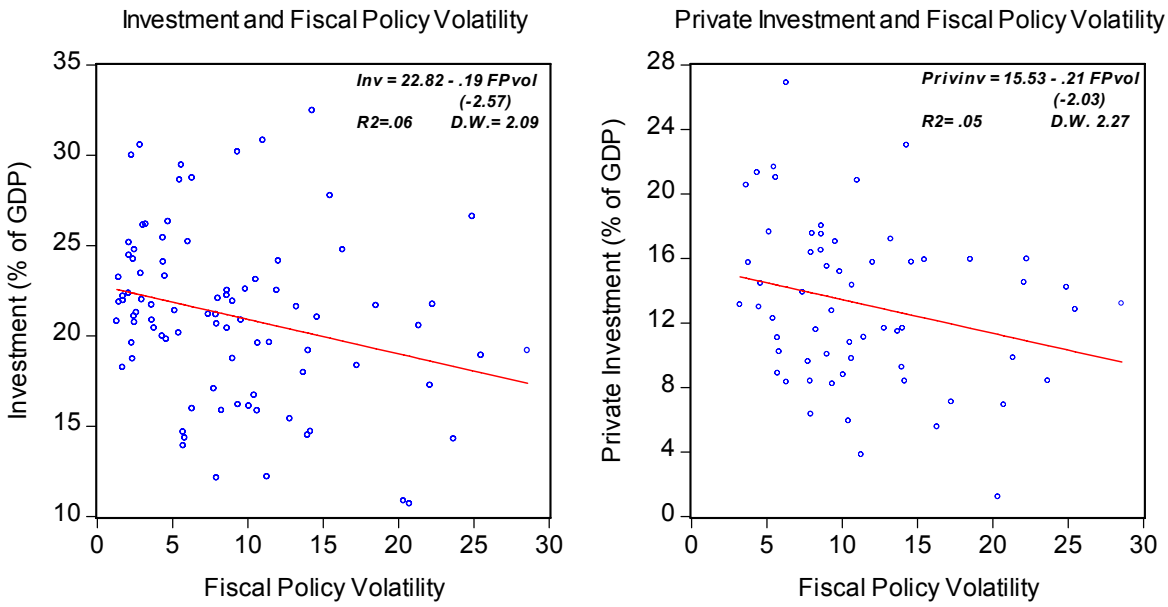
Before quantifying the welfare loss associated with consumption volatility, we discuss the more direct impact of policy volatility on growth and capital formation. Fiscal policy volatility is systematically associated with higher output volatility; which, in turn, is associated with lower growth.[3] This relationship is stronger in poorer countries. Researchers have verified econometrically this negative association, estimating an almost one-to-one relationship: a one standard deviation increase in volatility (.75 percentage points) is associated with lower growth by .75 percentage points. (Fatas and Mihov, 2003 and 2005) [4] The effect of volatility on growth is mostly through its effect on capital formation, but the relationship has to be interpreted with caution as there is heterogeneous response across countries: the negative impact of volatility on growth is stronger in poorer countries with more underdeveloped institutions, and has become stronger through time (Hnatkovska and Loayza, 2003).

[3] A survey of the research linking volatility with growth can be found in Loayza, et. al. (2007).

[4] Fatas and Mihov define the volatility of the discretionary component of public spending as the standard deviation of the residual of a regression of the growth rate of public consumption spending and the growth rate of output with other control variables including the country's income level and time trends. They use a sample of 91 countries over 1960-2000. The ranking of the countries by their volatility level using the Fatas-Mihov data (available in Fatas' homepage) is very similar to the ranking obtained this crude measure: the rank correlation coefficient is .85

Policy volatility adversely affects investment, as evidenced from micro and macro analysis. Firm-level data show that economic unpredictability and policy instability have a statistically significant adverse effect on expected sales growth (Chong and Gradstein, 2006)[5]: a one-level increase in policy instability decreases firm sales growth by almost 3 percent. At the country level, fiscal volatility is negatively associated with overall aggregate investment of the economy, confirming previous findings (Fatas-Mihov, 2003). When investment is disaggregated to contemplate private investment only (the sample is reduced due to data availability), the negative association persists as Figure 3 shows.

Figure 3



The "growth champions" have lower fiscal policy volatility, more efficient public sectors, and more centralized management of public finances. A recent survey of the growth process across the world in recent decades highlighted the experience of ten countries, labeling them the "growth champions": China, Korea, Thailand, Singapore, Vietnam, Hong Kong, Indonesia, Ireland, Chile and Malaysia (Harberger, 2005). Besides the commonalities analyzed by Harberger, namely the productivity increases and export performance, this group has other common traits related with fiscal policy: first, they have low volatility of fiscal policy; second, they have the highest scores of efficiency of the public sector (Afonso, Schuknecht, and Tanzi, 2006); and third, they have relatively centralized fiscal management, with the possible exception of China This section deals

[5] The World Bank Business Environment Survey asks firms to judge on a four point scale how problematic are the following four factors for operation and growth of their business: Policy instability (1) no obstacle; (2) minor obstacle; (3)moderate obstacle; (4) major obstacle.

with the first feature, namely the volatility of public spending, while the other two will be discussed in later sections.

Public expenditure volatility is lower in the "growth champions". Seven of the ten growth champions identified by Harberger are included in the Fatas-Mihov paper, and six of them rank below the mean indicator of policy volatility for developing countries. An alternative measure of volatility[6] shows that the "growth champions" have lower fiscal policy volatility than the average group of countries (Figure 2).

Policy volatility is directly associated with private consumption volatility. The link between the government's budget and household budgets results from the transfers made by the governments or the taxes paid by households. Hence, volatility in the government's finances is very likely to reflect in household finances, and hence on consumption decisions. Effectively, there is a clear positive association between fiscal policy instability and household consumption volatility. Simple linear regression indicates that almost 50 percent of the household consumption volatility is explained by public expenditure volatility (Figure 3).[7]

Consumers generally prefer smoother patterns of consumption, especially when it is close to subsistence levels, and they would be willing to sacrifice some amount of consumption to obtain a less risky consumption path. The amount that households are willing to sacrifice will depend on their tolerance to risk (degree of risk-aversion), and the amount of risk to which they are subject. The magnitude of consumption that households are wiling to forgo to obtain a more stable consumption level is the welfare cost of consumption volatility and reaches an astronomical 8% of consumption per year, on average, in the sample of developing countries considered in this paper (Table 2 and Appendix 1). For African countries the average is slightly higher, at 9%, while for Latin American countries it is 6%.

[6] To enlarge the sample and update the original Fatas-Mihov work , volatility was defined as the standard deviation of the growth rate of per capita public spending in each country in the period 1960-2005. The volatility ranking is very similar to that of Fatas and Mihov (the rank correlation between both orderings is .85), suggesting that the crude volatility measure we chose will not be biased by variables such as the income level. Ongoing work explores this possibility more rigorously.

[7] This figure is very similar to the Fatas-Mihov (2003) R^2 estimates for the relationship between output volatility and fiscal policy volatility. To capture the possibility of endogeneity of fiscal policy volatility, Fatas-Mihov use an instrumental variables method wich does not change the statistical significance of the simple OLS estimates, but does magnify the impact of policy volatility on output volatility.

Table 2
Costs of Output and Consumption Volatility*.

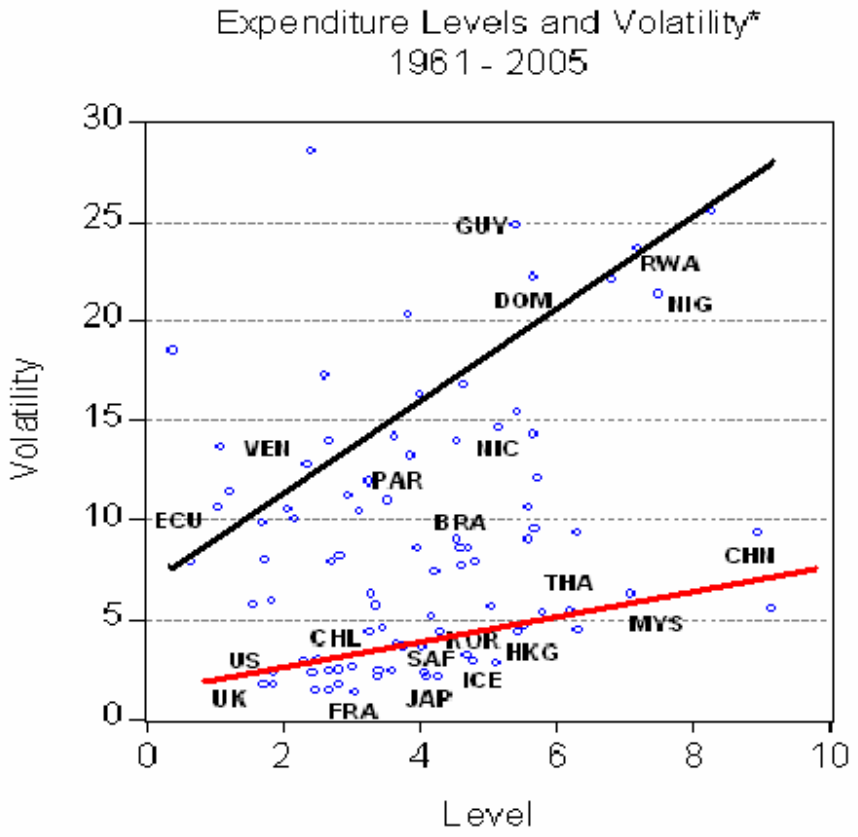
	Output	Consumption
Egypt	.004	.007
Lesotho	.041	.272
Mauritius	.000	.004
Morocco	.116	.160
South Africa	.027	.033
Tunisia	.011	.010
Uganda	.016	.044
Zimbawe	.154	.680
Argentina	.194	.244
Brazil	.013	.137
Chile	.036	.048
Colombia	.023	.055
Costa Rica	.026	.047
Ecuador	.039	.061
Mexico	.040	.069
Nicaragua	.022	.408
Venezuela	.211	.174
Australia	.001	.003
Canada	.012	.009
Denmark	.004	.012
Finland	.042	.037
France	.003	.004
Germany	.004	.005
Ireland	.030	.019
Italy	.001	.010
Japan	.004	.003
Spain	.003	.007
U.K.	.004	.008
United States	.004	.005
China	.011	.018
India	.015	.016
Indonesia	.095	.072
Korea	.054	.140
Russia	.230	.138
S. Arabia	.028	.027
Turkey	.106	.115

*percentage increase in output or consumption producing a welfare gain equivalent to that achieved by reducing volatility to the global average

Source: Author's calculations

For policy purposes, it is critical to determine the causes of fiscal volatility. Most studies conclude that the variables are related to political institutions and the constraints on the decision-maker (Fatas-Mihov, 2003, 2004 and 2005). However, Figure 3A highlights that higher spending is positively associated with higher volatility.[8] The positive association can be refined further by noting that there are two distinct groups of countries: the first, composed mostly of the OECD and East Asian economies, show a positive relationship between the variables, but volatility is significantly lower for any level of expenditure growth; Chile and South Africa are included in this group. The second group, composed mostly of Latin Amercan and African economies, also shows a positive association between the level and the volatility, though much steeper. In this case, higher expenditure is associated with more volatility, increasing at faster rate than in the other group of countries. The "growth champions" are all on the lower volatility schedule.

Figure 3A



*The level is the average growth rate of public spending and volatility is the standard deviation of the series.

[8] Policy is defined as as the growth rate of government spending per capita (1961-2005) while volatility refers to the standard deviation of the series. Gali (1994) presents a real business cycle (RBC) model in which larger government size (measured by the tax and spending ratios to GDP) is associated with higher output volatility.

If between 40 and 50 percent of the total welfare cost can be attributed to public expenditure volatility[9], the resulting cost has to be contrasted with the potential benefits of public spending to determine the overall benefit to society of public spending. Even further, knowledge of how this welfare cost is distributed across members of society is necessary because the supposed beneficiaries of public spending could be disproportionately bearing the loss induced by volatility. If risk tolerance decreases with wealth, as evidence suggests (Ogaki, et. al. 2001), then the poor will bear most of the cost of volatility. Additionally, to compensate for risk, the poor will self-insure by choosing activities with low rates of return but smaller variance. They will change their production and savings decisions to compensate for increasing risk (Morduch, 1995; Jalan et.al, 2001), in ways that reduce their income and, hence, perpetuate poverty. More analysis of how societies do risk-sharing along the lines of Townsend (1994) for India and Ogaki et.al (2001) for Pakistan, would be beneficial. For the moment, these estimates indicate an extremely high aggregate cost of volatility, requiring high benefits of spending to compensate for that welfare loss.

Hence, gauging the growth impact of additional spending, as well as other impact on human capital formation and welfare, is important. The rest of the paper is devoted to these challenges.

B. Public expenditure, outcomes and growth at the macro level: Where is the chain's weakest link?

This section discusses the assessment of the relationship between public expenditure and growth, intermediated by the effectiveness of public spending in affecting outcomes such as educational attainment, the number of paved road kilometers, or the kilowatts of electricity delivered to users. Recent literature surveys by the IMF (2005) and the World Bank (2007) exemplify the high ambiguity about the relationship between public expenditure and growth at the macroeconomic level. On one hand, the IMF survey reports that "empirical evidence on the impact of public investment on growth remains mixed" and that "their [infrastructure projects] impact on GDP growth is more uncertain. Empirical studies that have tried to estimate such impact have yielded widely different results…. Empirical ambiguity over links between spending and growth has also been shown to exist for other types of expenditure..." (IMF 2005, pg. 13). On the other hand, the World Bank survey states that "…though some non-robustness remains in the evidence, this largely applies to studies before 2000. … a number of recent papers reverse the earlier findings, and now find strong positive growth effects from public capital spending. [list of papers] show that sectors usually regarded as productive (transport and

[9] Based on the percentage of the variance of household consumption volatility explained by public expenditure volatility (R^2 of the regressions presented in figures discussed in the text and footnote 5). This assumes that policy volatility is exogenous with respect to consumption volatility, based on the instrumental variables estimation of Fatas-Mihov (2003) referring to output volatility. This work is yet to be done for consumption volatility, but similarity of the country rankings by volatility between the Fatas-Mihov papers and this paper's, suggests results will not be radically different.

communications, education, health, etc.) have significant positive long-run growth effects…" (WB, 2007 pg. 42-43)

The ambiguity of results is not surprising, as the relationship between public expenditure and growth is mediated by a long chain subject to multiple weak links that are not necessarily the same across countries, or even across expenditure programs within the same country.[10] This section examines the links between public expenditure and capital formation with particular attention to three topics: 1) the need of considering the link between expenditure and revenue, given that the additional funding will have an impact on growth; 2) the necessity of differentiating spending flows (investment) and capital stocks; 3) the importance of recognizing heterogeneity in output responses across countries, across types of infrastructure, and across levels of network development. The above considerations will require a more detailed analysis but will provide a more useful diagnosis of the role of spending on growth.

1. The impact of additional funding to finance more spending

Regardless of the productivity of expenditure, both the policymaker and the economic analyst must consider the budget constraint, that is, how the additional expenditure will be funded. It is essential to include in the analysis of public investment the cost to society of raising an additional dollar to pay for more spending. This will be discussed in the next section, and momentarily, the focus is the distorting effect of the tax system and its negative growth effect. The impact of productive expenditure on growth will depend on the relative productivity of the private sector vis. a vis the public sector, on the size of the productive expenditure, and on the size of public sector. Both the size of public spending and public debt matter because they imply higher taxation levels which entail increasing more than proportionately the marginal cost of public funds. Appendix 2 summarizes a simple analytical model that shows how, unless public expenditure is extremely productive relative to the private sector and public debt is quite low, increasing taxation to fund productive expenditures will have a negative growth impact.

From the econometric viewpoint, the above consideration implies imposing a restriction when estimating the impact of expenditure on output. Otherwise regression models would be mis-specified (Kneller, Bleaney and Gemmell, 1999; Bleaney, Gemmell, and Kneller, 2001). Recently, many papers have estimated the growth impact of different types of spending (productive and unproductive) jointly with alternative forms taxation (income taxes vs. consumption taxes). Two good recent examples can be found in Adam and Bevan (2005) and World Bank (2007).

These studies have several common features and results: 1) they are panel studies, though covering different samples;[11] 2) both studies consider public investment (instead of public capital) in a category labeled "productive expenditure" that also includes health

[10] The analogy of the chain with weak links is original from Filmer, et.al (2000) that analyze the relationship between public spending and health outcomes.

[11] The Adam-Bevan paper includes 45 non-OECD countries, while the World Bank sample is limited to 21 countries of eastern Europe and Central Asia;

and education spending; 3) both papers show that an increase in unproductive expenditures is associated with lower GDP growth rates, while productive expenditures are marginally significant with a positive growth effect; 4) the impact of taxes is unambiguously negative and significant[12], with the magnitude of the coefficient similar to the elasticity of output with respect to productive expenditure.

For policy purposes the main lesson is to consider jointly the effects of both spending and taxation on economic growth. The signs and magnitudes of the tax and expenditure variables in both studies cited above, indicate that both effects offset each other when dealing with productive expenditures, but in the case of unproductive expenditures the effect will be unambiguously negative. These results explain why many countries that instituted generous pension benefits in the past, and have had to reduce productive spending or raise current taxes to pay those liabilities, experience slower growth, such as Brazil, Uruguay, and Colombia.

2. The difference between investment (flow) and public capital (stock)

Besides considering the effect of the additional resources required to fund spending, policymakers must bear in mind that growth and welfare depend on the flow of services provided by capital, rather than on spending itself. The services, in turn, are better approximated by the capital stock, while investment spending is associated with changes in it. Hence, the amount of investment that actually is transformed into public capital, or the efficacy of investment, is an essential element to incorporate in the measurement of the impact of public spending on growth or estimations of the rate of return of public spending. Initially, in the area of infrastructure, researchers were limited by data availability on capital stocks and public investment. Hence, researchers inferred the efficacy of public investment from aggregate macro data and results indicated that between 40 and 50 percent of investment was incorporated into physical capital, with great variability across countries (Pritchett, 1996 and 2000).[13]

These findings have clear analytical and policy implications. From the analytical viewpoint, the ineffectiveness would explain why some studies don't find a relationship between public investment and growth. When simulating the growth impact of changes in public capital spending it would be advisable to include some measure of the effectiveness of investment (Agenor, et. al. 2007). From the policy standpoint, changes in public capital spending are not linked one-to-one with changes in output. In terms of

[12] In the Adam-Bevan paper, while in the WB paper the tax coefficient is negative and significant in three out of 4 cases

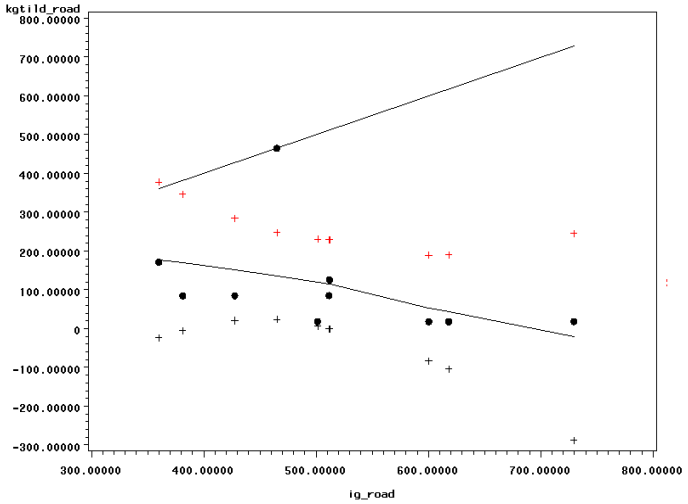
[13] Pritchett estimated the efficacy of investment based on a growth accounting exercise. He estimated the rate of capital accumulation that would be consistent with a "true" Total Factor Productivity (TFP) growth of between 0 and 1 % per year. This "implicit" rate of factor accumulation would be compared with the observed rate of capital accumulation. Generally the implicit rate would be much smaller than the observed rate, leading Pritchett to conjecture that recorded (actual) figures overestimated the real capital accumulation due to the low efficacy of investment. Efficacy fluctuated between 8% and 95%, depending on the TFP benchmark (0 or 1%) and the region, with Sub-Saharan Africa and South Asia exhibiting the lowest efficacy and the high performing Asian economies reaching (95%). The OECD average fluctuated between 74% and 98%.

policy advice, it would be inappropriate to conclude that a specific country increase public spending in roads because a cross-country panel regression coefficient is high, or to conclude that the country should cut spending because the regression coefficient is low or non-significant.

More recent studies, based on new datasets and more sophisticated methods, estimate the efficacy of public investment in the range between 40 and 50 percent. Data on infrastructure stocks across countries (Canning 1998, 1999) and of public investment in infrastructure (Calderon and Serven, 2003) facilitated estimation of public capital stock series. Some researchers opted for considering the accumulation of past investment as a proxy for changes in public capital, while others correlated the information in both datasets. A complete and rigorous study shows that, in developing countries, the perpetual inventory method (PIM) is a poor approximation of the function that relates public investment and physical capital: estimations indicate that only 38 and 40 cents of every investment dollar is transformed into public capital. For instance, Figure 4 shows the estimated relationship between investment and capital stocks in the roads sector in Colombia, and Figure 5 depicts the same relationship for overall investment in Mexico. Both cases show the tenuous relationship between one and the other. In developed countries the same approximation is almost identical to what the PIM would produce, reflecting high efficacy of investment. Figure 6 shows the close approximation between investment in roads and (changes in) capital stocks in the United States (Arestoff and Hurlin, 2006).

Figure 4

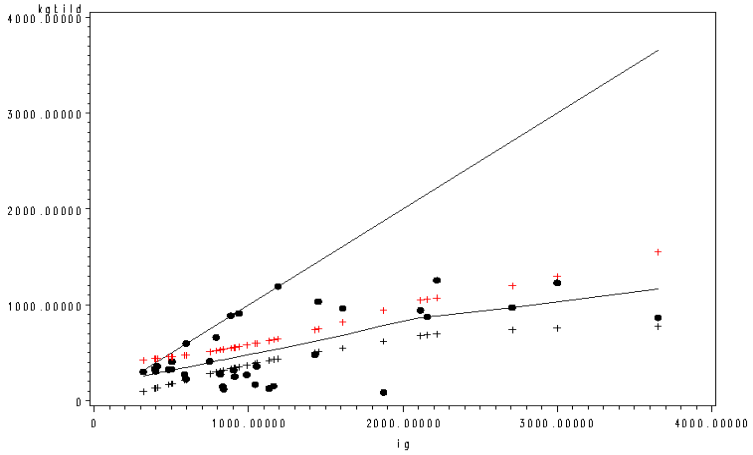
Non-Parametric Estimated Efficiency Functions of Sectorial Public Investment. Colombia, 1980-1994 (US$ million, current prices)
Roads



Source; Arestoff and Hurlin 2006

Figure 5

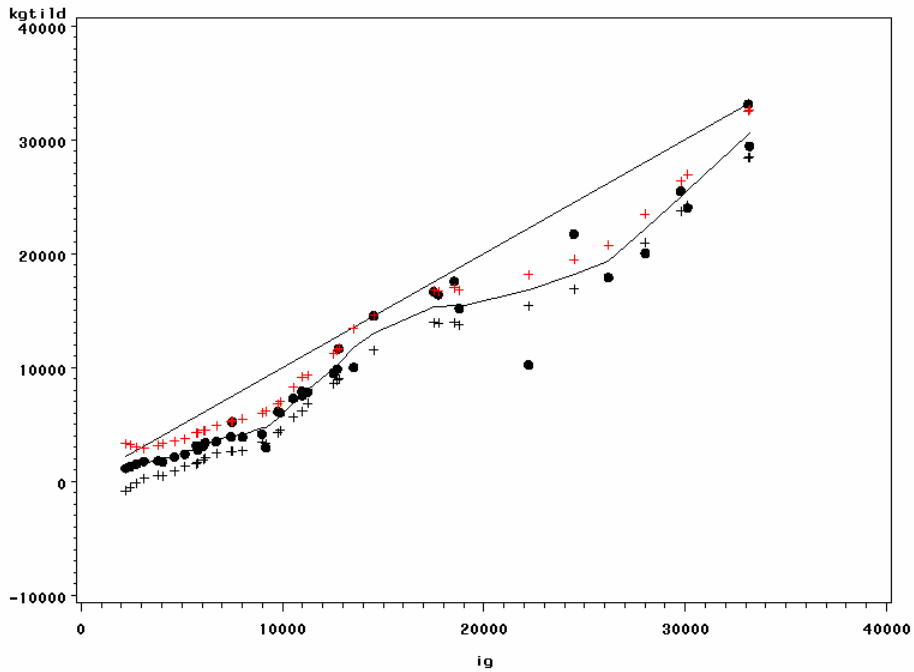
Non-Parametric Estimated Efficiency Function of Total
Public Investment. Mexico, 1980-1994 (US$ million, current prices)



Source; Arestoff and Hurlin 2006

Figure 6

Non-Parametric Estimated Efficiency Function of Public Investment
in Streets and Highways . United States, 1951-1992 (US$ million, Historical Cost)



Source; Arestoff and Hurlin 2006

The efficacy of public spending is also a challenge in the formation of human capital as evidenced in education, nutrition, and health programs. In education, the resources that reach the student are not necessarily the same that leave the central budget office. For

instance, public schools in Uganda received only 13 percent of the central government's allocation for non-wage expenditures, with the leakage depending on the size of the school and the wealth of the families in the particular district. When the scool transfers were published in local newspapers and schools were required to publish the inflow of funds, the leakage reduced substantially to 10 percent (Reinnika and Svensson, 2001). In nutrition, programs targeted to ensure children's intake of minimum nutritional elements are also subject to leaks. In Peru, the *Vaso de Leche (Glass of Milk)* program oriented to ensure children's intake of calcium and protein was considered a success until a World Bank Public Expenditure Tracking Survey (World Bank, 2002) found that the beneficiaries (children under six) received 30 cents for every dollar spent by the central government. Private gains and high operational costs of the public sector explained the leakage of funds.

The health sector also registers weak links between public spending and health outcomes that have been explored comprehensively across numerous counties (Filmer, et.al., 2000). The pioneering study in this area showed that inadequate institutional capacity and allocation of budgets to the activities in which market failures are not that big limit the potential impact of public spending on improving health outcomes. The authors showed that, when conditions were not serious or quality was not important, patients would choose public facilities. But for serious conditions, individuals were willing to pay for private higher-quality care, bypassing the public service provider.

Improving service delivery for human capital formation is more difficult than in physical capital because outcomes are observable only with long lags and it is difficult to monitor the work effort by service providers. This limits the option of linking pay to performance and suggests that, in many circumstances, to improve service delivery policymakers will have to rely more on individual choice in a setting involving a proper mix of public and private provision of services. Generalization to a uniform standard of private sector participation is impossible, as there are a constellation of possible arrangements depending on whether the service is homogeneous or not, on the communities preferences, and on the monitoring feasibility. In general, the more heterogeneous is the service, and the higher the difficulty in monitoring outcomes and work effort, the more involved will local governments have to be. If the service is more homogeneous, the central government may be involved. In both cases, the extent of private provision will depend on monitoring capabilities (World Bank, 2003)

3. Is all capital equally productive?

1. Human capital

The output response to additions of human or physical capital will vary with its quality and productivity. This heterogeneity helps explain why the macroeconomic literature based on aggregate figures finds ambiguous results on the impact of public investment or capital on growth. This section examines the sources of heterogeneity in human and physical capital, and discusses its implications for policy.

Measures of human capital in most growth regressions, including literacy rates, enrollment ratios or the average years of schooling, are unreliable proxies for human capital because they do not incorporate the rate of return to education nor the quality of education. The most commonly used proxy for human capital is the number of years of schooling. The simple aggregation ignores that rates of return vary with the education level and that schooling systems may deliver different skill levels in the same time period, depending on their quality When human capital is adjusted to include these two effects, results consistently confirm the crucial role of human capital in explaining development differences across countries: cross-country differences in the stock of human capital account for one half of the dispersion in economic development level across the world, and for practically all the differences across the OECD countries (WoBmann, 2003). However, the measure of human capital depends critically on varying rates of return to education and quality of education adjustments.[14]

Recent estimates of the rate of return to education fluctuate around 10 percent, though dispersion across regions and countries is large. While the average rate in developing countries is 11 percent, in the OECD economies it is only 7.5 percent (Psacharopolus and Patrinos, 2004: Patrinos, et.al. 2006). Rates of return also vary for primary, secondary and higher education, showing a decreasing trend as the level progresses from primary to more advanced levels. Social rates of return also vary with the country's income level: for instance, primary education rates of return in low income countries average 21%, in middle income countries around 19%, and in high income counties about 13%. Rates of return also vary across income groups within the same country: returns to education for the bottom quantile of the income distribution are 3 percentage points higher than for the top one in a sample of East Asian countries, while in Latin America they are 5 percentage points lower (Patrinos, et.al, 2006).

The variability of rates of return suggests potential pitfalls of working with country average rates and might explain the contradiction between the micro literature showing consistent returns of education with the earlier macro literature showing almost no relationship between education expansions and growth. Pritchett (2001, 2006) argued that deficiencies in the institutional setting accounted for cognitive skills being directed to unproductive activities, rendering negligible average effects of education on growth. This hypothesis is verified in growth regressions that include interaction terms between educational variables and institutional proxies such as openness of the economy or indicators of the rule of law (Hanushek and WoBmann, 2007).

Another critical element in measuring human capital is the quality of education. Growth regressions that include both the typical education quantity indicators (years of

[14] Dessus (2001) estimated a neoclassical growth model with a heterogeneous elasticity of output with respect to human capital across countries. The elasticity was a function of several variables that approximated the quality of education system, including the school infrastructure, the inequality in school achievement across different groups of society, and the pupil/student ratio. Results indicate that human capital is an important factor explaining cross-country differences, once heterogeneity in the school quality was allowed to operate. Hannushek and WoBmann (2007) also consider inequality in the distribution of education quality across income groups reaching similar conclusions.

schooling) and the more novel quality of education indicators[15] substantially improve the explanatory power of the model. Additionally, the quantity indicators turn out statistically insignificant (Hanushek and WoBmann, 2007). Recognizing the limitations in the use of standardized testing as proxies for the quality of education, it is essential to explain how to achieve improvements in the quality of education. This leads us to public expenditure and institutional issues.

The literature, in general, shows that additional public spending is not necessarily associated with improved education outcomes (Pritchett 2001, 2006; Hanushek and Kimko, 2000). Hanushek and WoBmann (20007) corroborate these results using the quality of education indicators for a sample of OECD and non-OECD economies. The authors verify graphically and with linear regression the absence of a relationship between educational spending per student and student performance measured as the math test scores in the PISA 2003. Though these results were robust for both the average OECD country and the non-OECD, it is legitimate to ask whether they are valid for an enlarged sample including more developing countries and whether policy design should be based on average behavior.

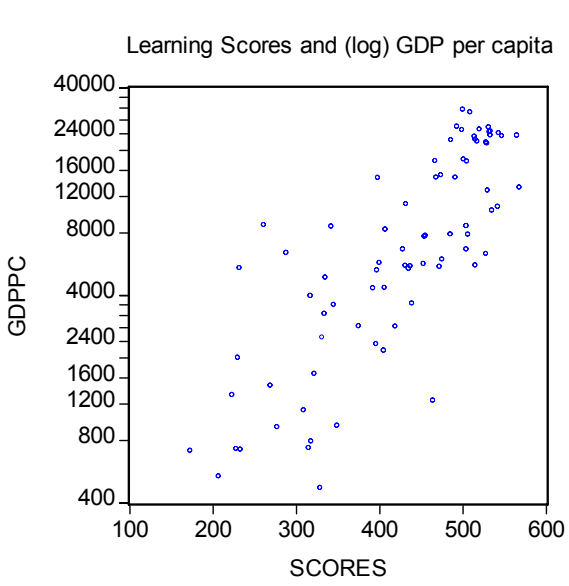
When the sample is enlarged to include more developing countries, the association between public spending and quality of education for the average country remains insignificant and may even turn negative. The comparison across countries of such a diverse nature deserves caution. For instance, there is a clear positive association between the GDP per capita level and the test scores, as Figure 7 shows. Similarly, there is a positive association between public spending in education and GDP. When both scores and educational spending are controlled for GDP, the negative association emerges for the average country as Figure 8 shows.

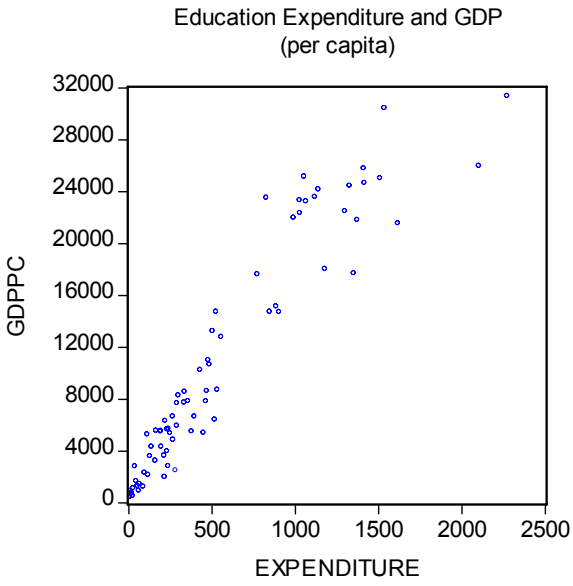
However, for policymaking purposes, working with the best-practice cases may be more informative and useful than working with average behavior. This is especially relevant when adopting a production-function approach to education that relates inputs (spending) and outcomes (education attainment). In this case, what is of interest is the maximum expected output for a given input level, or alternatively, the minimum amount of input required to achieve a determined output level. The countries that obtain the highest scores per unit of public spending (best-practice cases) are considered the most efficient and constitute what is known as the efficiency frontier, depicted in Figure 9. The other countries' efficiency is estimated as the distance with respect to the frontier. When the relationship between public expenditure and quality of education is cast in these terms, an efficiency frontier emerges. Instead of a negative association, a step-wise linear function is obtained with the best-practice (most efficient) cases lying on the frontier.[16]

[15] The indicator is based on results form cognitive achievement tests, such as the TIMSS, the PISA and the PIRLS among others.

[16] This method is called the Free Disposable Hull (FDH) method. A similar technique, Data Envelopment Analysis (DEA) can also be used to examine efficiency. Appendix 3 presents a discussion of efficiency frontiers.

Figure 7

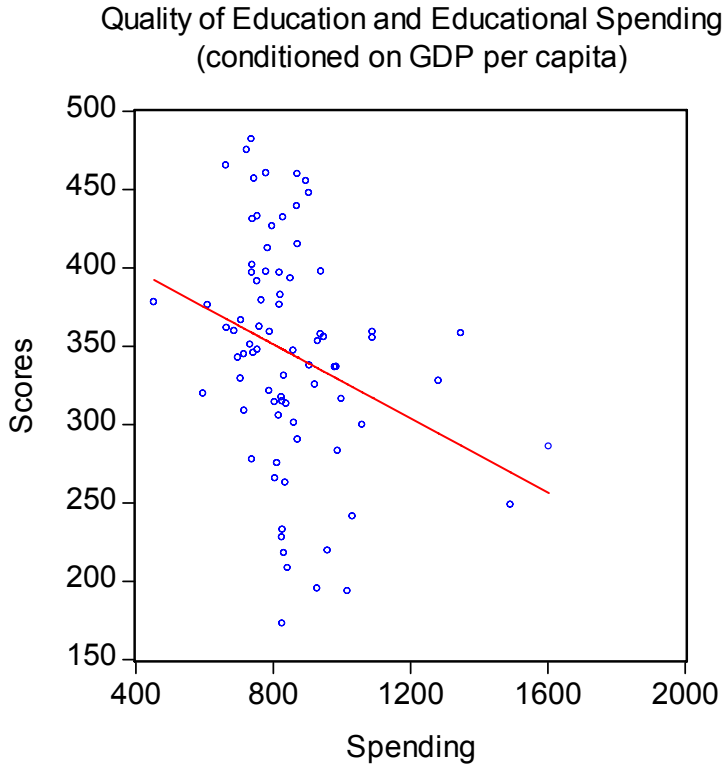




source: Herrera and Pang (2006)

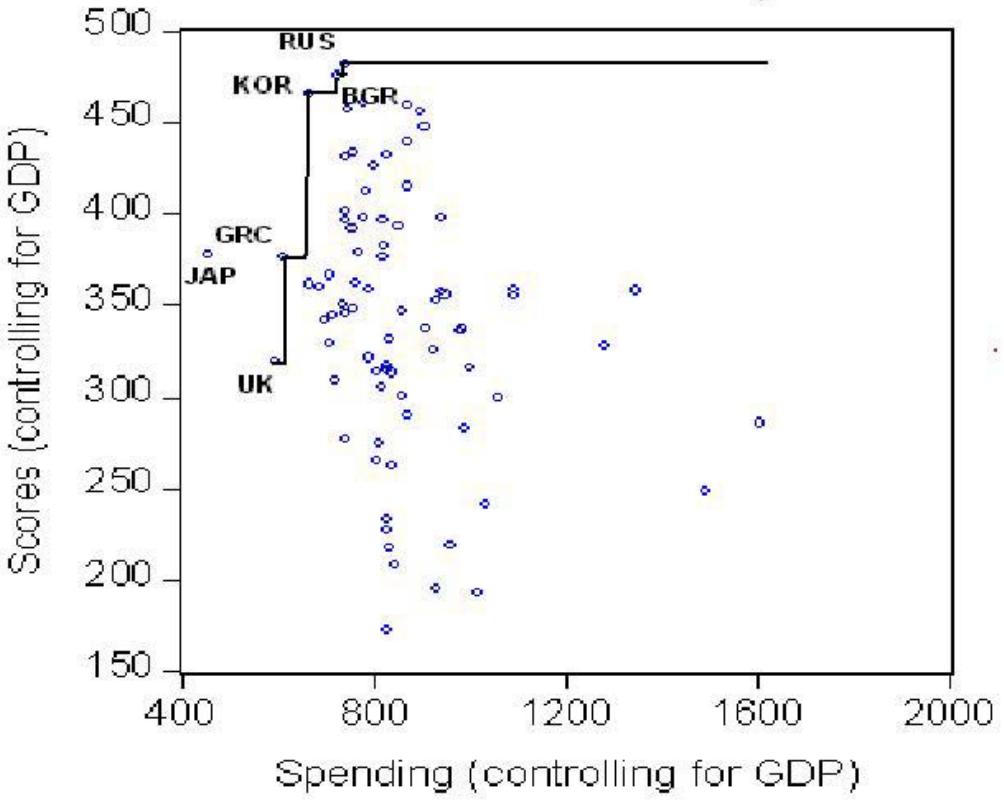
.

Figure 8



Source: Herrera and Pang (2006)

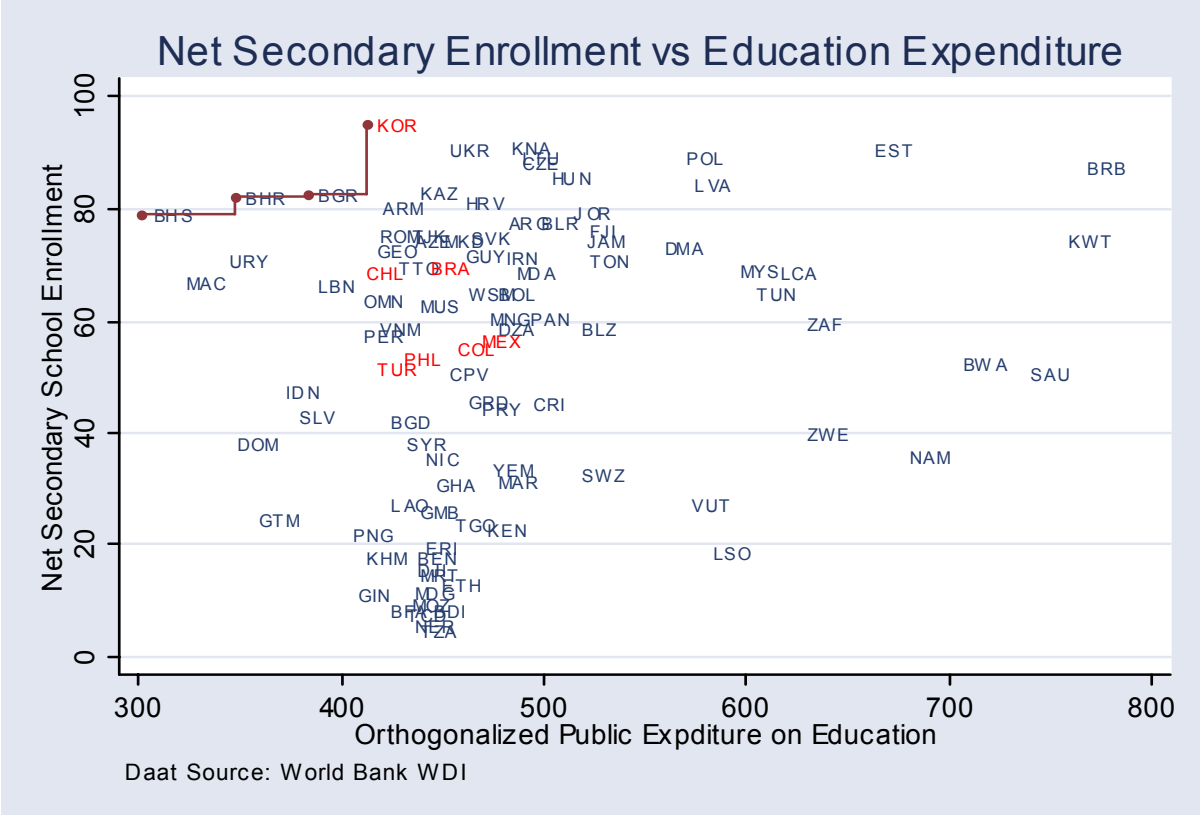
Figure 9



Source: Herrera and Pang (2006)

These empirical measures of efficiency, based on ratios of observed output levels to the maximum that could have been obtained given the inputs utilized date to the 1950s (Debreu, 1951; Farell, 1957) and have recently been used to estimate the efficiency of public spending on health and education (Gupta and Verhoeven, 2001; Afonso et al. 2004; Herrera and Pang, 2006) The focus of the recent work in these two sectors is due to their relative importance within public budgets and to the availability of cross-country data to perform the benchmarking. These studies use aggregate spending on health or education as an input, jointly with other inputs such as the adult literacy rates, the ratio of teachers to pupils, or the amount of private spending, to obtain efficiency frontiers for education or health. The efficiency frontiers estimated in this fashion are used to benchmark the efficiency levels of the other countries. For instance, in the case of secondary enrollment, Bulgaria and Korea appear on the frontier ((Figure 9A), while countries like India and Uganda show that, given their expenditure levels, secondary enrollment rates are 47 percent and 13 percent of the most efficient countries. Peru scores 78 percent. (Herrera and Pang, 2006).

Figure 9A



Source: Herrera and Pang (2006)

In terms of policy implications, it is vital to differentiate between the technically efficient level and the optimal or desired spending level. Even if a country is identified as an "efficient" benchmark country, it may very well still need to expand its public spending levels to achieve a target level of educational or health attainment indicators. Such is the case of countries with low spending levels and low attainment indicators, close to the origin of the efficient frontier. The important thing is that countries expand their scale of operation along the efficient frontier.

This measure of technical efficiency, unfortunately, has nothing to say about the optimal spending level. It simply states benchmarks of costs per unit of output, focusing on doing things right (at the lowest possible technical cost) but an important consideration is if it is doing the right things. This benchmarking method necessarily has to give way to analysis with more economical content Hence the discussion has to move to estimation of social benefits and costs, as well as causal inference, topics covered in the next sections.

The above discussion illustrates the point that policies that simply provide more resources will not necessarily improve education outcomes; it is imperative to examine other links of the chain. Teachers are key inputs in the production function and account for most of the cost of education. In both developed and developing nations teachers' unions have

contractual arrangements that are not conducive to monitoring performance. Though there is some evidence that performance-related pay leads to improving achievement (Atkinson, 2004; Vegas and Umansky, 2005, Duflo, et.al. 2005), it is not fully understood what constitutes high-quality teaching plus the difficulty in monitoring behavior might lead to improper behavior, such as the cheating documented by Jacob and Levitt (2003) in the Chicago school system. Other behavior frequently observed in developing nations is that of teacher absenteeism. Countries where absenteeism is more prevalent (Table 3), such as India and Uganda, ranked lower in terms of efficiency scores estimated in educational attainment, while those with lower absenteeism, like Peru, ranked higher (Herrera and Pang, 2006).

Table 3
Teacher Absenteeism

	Percentage of teachers that were absent this many times in 2 visits (3 in India)			
	0	1	2	3
Bangladesh	73.4	23.5	3.2	
Ecuador	82.8	6.9	10.4	
India	49.1	32.7	13.5	4.8
Indonesia	67.7	27.5	4.8	
Peru	81.0	17.3	1.7	
Uganda	63.0	29.6	7.4	

Note: Interpretation of the figures. 73.4 % of the teachers in Bangladesh were never absent; 23.5 % were absent once, and 3.2 percent were absent twice.
Source: Chaudhury, et.al.

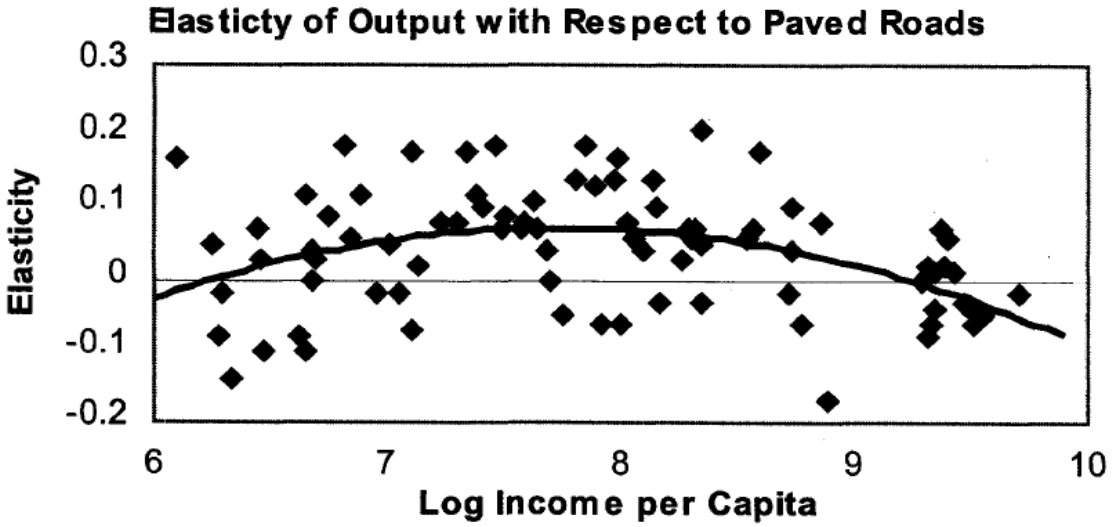
Without an adequate institutional setting, educational spending will lead to negligible changes in human capital and hence on growth. A package of institutions or policies that would provide this institutional setting would: 1) promote more competition so that parental demand generates incentives to individual schools; 2) allow autonomy in local decision makers so that school mangers promote student achievement; and 3) develop an accountability system so that rewards are based on school performance (Hanushek and WoBmann, 2007).

2. Physical capital

In the case of physical capital (infrastructure), studies show that its growth impact depends on its quality and is heterogeneous owing to multiple factors: different types of physical capital, differences in the stage of network development, differences through time (Calderon and Serven, 2004; Hurlin, 2006; Colletaz and Hurlin, 2006). Different types of capital will have distinct effects on output. For instance, roads and railways will not have necessarily the same impact on output. Similarly, a single road with no connections will not have the same impact as one interconnected with the main highway system. These differences in the determinants of capital productivity are seldom

incorporated in econometric exercises. Initially, to allow the possibility of differing productivities across countries, researchers split their samples in poor and rich counties, with a resulting small and insignificant elasticity of output for the lower income group (Canning and Benathan, 2000). That outcome was obtained with a simple (and restrictive) Cobb-Douglas production function that imposes a constant and homogeneous (per each group) elasticity of output with respect to public capital. Using a more flexible functional form (translog), that captured non-linearity by introducing quadratic terms and interaction between variables, the authors obtained elasticities that varied with the level of income (Canning and Benathan, 2000)

Figure 10.



Source: Canning and Benathan (2000)

The non-linearity and heterogeneity in the response of output to physical capital may arise from sources different than the countries' income level. The countries' relative intensity of factor utilization (capital per worker) may be the differentiating variable. In particular, the stage of network development may be the source of heterogeneity in the output response to additional public capital. If this were the case, the arbitrary classification of countries into predetermined categories imposes (incorrect) restrictions in the estimation. A more flexible approach would allow endogenous classification of countries into groups according to the stage of network development, with possible switching from one group to another when endogenously determined threshold levels are achieved. This approach has been applied to both samples of OECD countries (Colletaz and Hurlin, 2006) and developing economies (Hurlin, 2006). The main implication is that productivity is allowed to vary by type of capital, by the particular stage of network development, and through time.

This flexible approach applied to a sample of OECD economies yields interesting results, with the advantages that it facilitates comparison with well-known results and the quality of the data may be more reliable (Colletaz and Hurlin, 2006). Using individual country

time series (ignoring the panel dimension), the authors obtained very high elasticities of output with respect to capital: .39 for the United States, identical to Aschauer's (1989) pioneering result. Australia, France and Germany also register values of .4 or above. These estimates imply rates of return of public capital of over 100%.[17] By now, the literature recognizes that this result is explained by multicolinearity and reverse causation.

These limitations were overcome by exploiting the panel dimension and using instrumental variables or GMM methods, leading to lower estimates of the productivity, but still implying rates of return higher than economic analysis of projects would indicate reasonable (discussed in the following section). However, when the estimation method allows for heterogeneity across countries depending on the level of infrastructure network development, the estimated elasticity falls to between 6 and 9 percent in the US, France, Germany, Canada and the UK while Italy and Australia are in the 10 to 14 percent range.

The non-linear effect of infrastructure on output is verified when the sample is extended to include developing countries (Hurlin, 2006). In addition, by differentiating between types of capital, i.e. roads, electricity, telephones and railways, much more flexibility is added. However, it becomes more difficult to specify a value for the productivity of public capital. For roads, the productivity of capital in early stages of development is around 6 percent, increases to 8 percent in the next stage of development, and then decreases until reaching 3 percent in the most mature stage. In electricity, the productivity parameter starts quite low, at 4 percent, increases quickly to 10 percent and remains about that level. In telephone, the elasticity begins at a very high 18 percent and decreases very gradually until about 16 percent in the last stage. The elasticity of output with respect to railway development starts at a low3 percent and increases slowly to 5 percent.

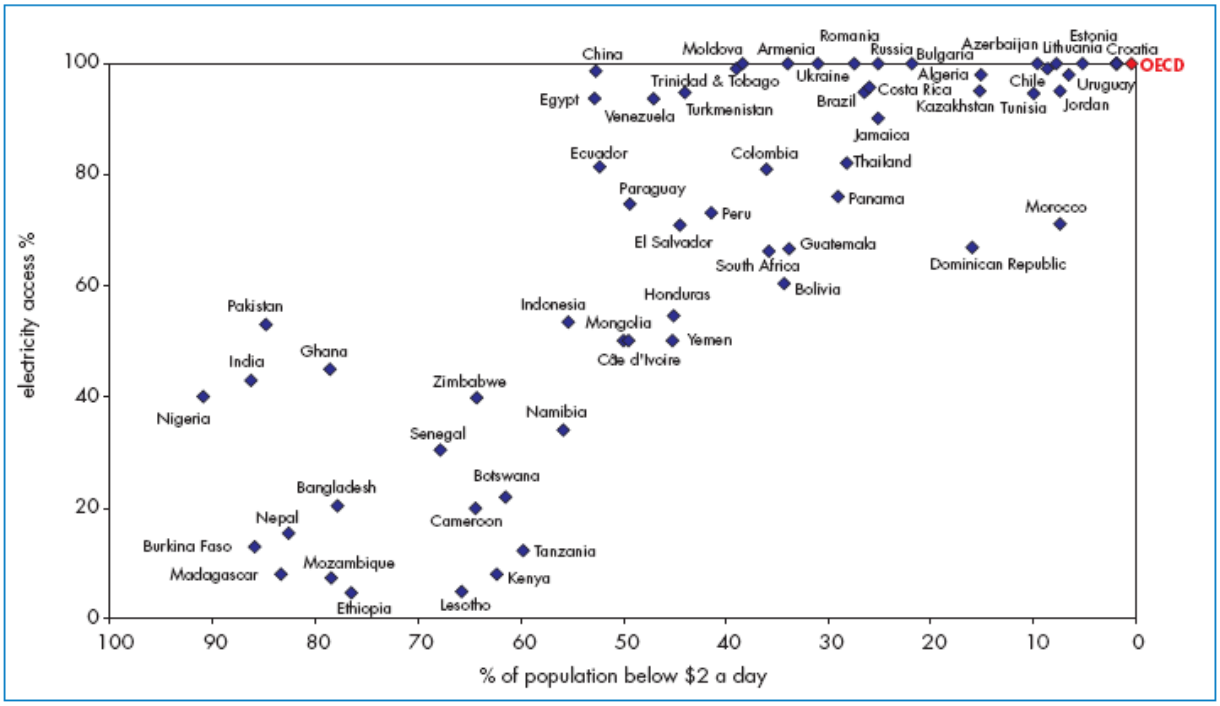
The critical lesson from the verification of the network effects is that establishing priorities for the capital spending program depending on the level of network completion is critical to maximize the growth impact of a given expenditure level. Additionally, though, historically, increases in particular infrastructure development could have been associated with growth spurts, those episodes may be unrepeatable if the network reached full completion and the productivity of additional capital is close to zero. As an example, consider the growth boost sparked by the building of the interstate highway system in the US during the 1950's and 1960's that is deemed unrepeatable because of the fall in the productivity in road infrastructure (Fernald, 1999).

Considering heterogeneity in the efficiency of resource use goes beyond the econometric implications discussed above, as there are other revealing comparisons in the efficiency and effectiveness with which countries use the resources. Even tasks that apparently

[17] Gramlich (1994) showed this with a modified production function expressing output (Q) as a function of private capital (K), labor (L) and government capital (G). If Q=A F (K,L, G), then the elasticity of output with respect to public capital, e, would be equal to $F_G G/Q$, where F_G is the marginal product of the factor, or the rate of return. With G/Q estimated around 40%, the elasticity of.4 yields a rate of return of around 100%.

depend on simple engineering principles, such as water and electricity distribution, reveal great heterogeneity in the efficiency and effectiveness in the use of resources. Given the association between indicators of coverage of infrastructure services and poverty (Fig. 11), it is essential to increase access to electricity and water across the world. In the case of electricity, both coverage levels and progress have been uneven across regions (Table 4), and, in general, there is much to be achieved.

Figure 11
Electricity Access and Poverty Rates across the World



Source: IEA, World Energy Outlook, 2002.

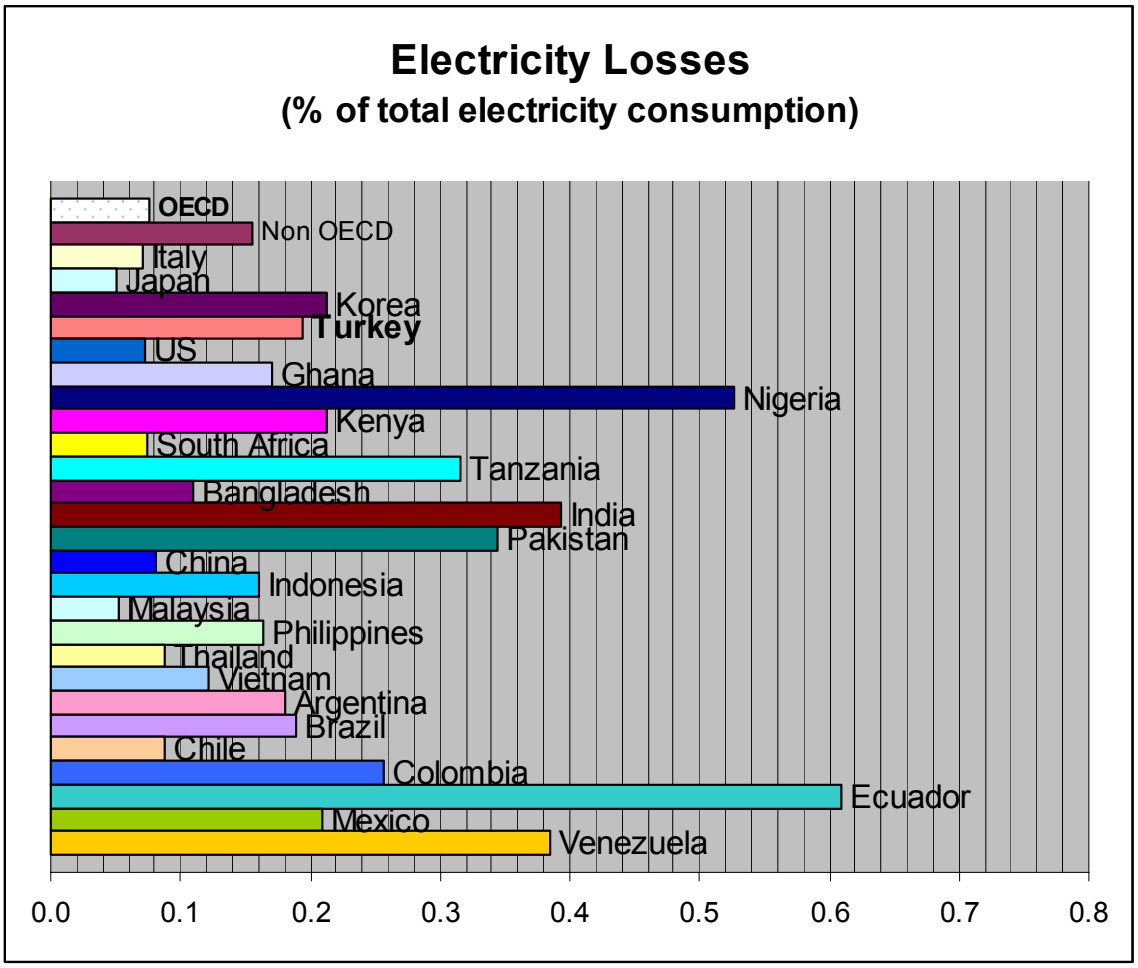
Table 4
Electrification Rates by Region, 1970-2005

	1970	1990	2005
North Africa	34	61	96
Sub-Sahara Africa	9	16	26
South Asia	17	32	52
East Asia/China	30	56	89
Latin America	45	70	90
Middle East	36	64	78
Developing Economies	25	46	68
Transition Economies and OECD			100
World	**49**	**60**	**76**

Source: IEA, World Energy Outlook, 2002, 2006.

Though part of the solution resides in capacity expansion, an integral approach must include reduction of transmission and distribution losses. These are different across countries and regions. In Africa, where the coverage ratio is the lowest, the losses reach astronomical levels, particularly in Congo, where the losses are equivalent to the full value of consumption, and Nigeria where they reach 53%, as shown in Table 5 and Figure 12. Other cases of large losses are Venezuela (38%) and Ecuador (61%). South Asia also shows extremely low access to electricity and extraordinarily high loss rates.

Figure 12



Source: Author's calculations based on IEA data, Table 5.

Table 5 Electricity Losses (% of electricity consumption)	
OECD	0.076
Non-OECD	0.155
Italy	0.071
Japan	0.051
Korea	0.213
Poland	0.144
Portugal	0.092
Spain	0.104
Turkey	0.194
UK	0.094
US	0.073
Angola	0.167
Congo	1.000
Ghana	0.170
Nigeria	0.527
Kenya	0.213
Mozambique	0.171
Senegal	0.158
South Africa	0.074
Tanzania	0.316
Zimbawe	0.149
Egypt	0.145
Argentina	0.181
Bolivia	0.128
Brazil	0.188
Chile	0.088
Colombia	0.256
Ecuador	0.609
Guatemala	0.048
Mexico	0.209
Venezuela	0.384
Bangladesh	0.110
India	0.393
Pakistan	0.344
China	0.081
Indonesia	0.161
Malaysia	0.053
Philippines	0.163
Thailand	0.087
Vietnam	0.121
Source: calculations based on IEA data	

Some countries expanded electricity generating capacity with little progress in increasing the access of the population to electricity, while others that increased capacity moderately but had low losses were able to expand coverage. Within the African continent there are striking contrasts. For example, Tanzania expanded generating capacity at an annual average rate of 8% during 2000-2004, but access to electricity remained constant (Table 6). Tanzania's loss rate is 32%. Ghana increased its installed capacity at an annual average growth rate of 22 %, while it coverage rose modestly by 4 percentage points. On the other extreme, South Africa's generating capacity rose by 1 percent, but coverage increased the same 4 percentage points, thanks to the low loss rates. India and Pakistan present an interesting contrast. While both have similar low loss rates and capacity expanded at the same rate, India increased coverage from 43% of the population to 56%, while Pakistan's electrification rates remained practically constant.

Table 6
Where did the Electricity Go? Electrification Rates and Percentage Annual Increase in Installed Capacity in Selected Countries, 2000-2005

	Coverage 2000 (%)	Coverage 2005 (%)	Average annual percentage change in installed capacity
Ghana	45	49	22
Kenya	8	14	13
South Africa	66	70	1
Tanzania	11	11	8
India	43	56	4
Pakistan	53	54	4
Indonesia	53	54	2
Thailand	82	99	8
Vietnam	76	84	11
Guatemala	67	79	17

Source: IEA, WEO 2002, 2006.

It is also imperative to reduce losses in the water sector, where utilities in developing countries register physical and commercial losses equivalent to 35% of the total input; in developed nations the loss ratio is only 15% (Table 6A). The monetary cost of these losses will depend on the marginal cost of water and the average tariff charged, but conservative estimates of the loss suggest that it is around US$14 billion. The magnitude of the estimate is appreciated when compared to the estimated US$20 billion required to invest every year to reach the MDG for basic access to potable water in developing countries (Kingdom, et.al. 2006).

Table 6A- Estimates of non-revenue water (NRW) across the world

			ESTIMATES OF NRW					
				Ratio		Volume (billions of m³/year)		
	Supplied population (millions, 2002)	System input l/capita/day	Level of NRW (% of system input)	Physical losses (%)	Commercial losses (%)	Physical losses	Commercial losses	Total NRW
Developed countries	744.8	300	15	80	20	9.8	2.4	12.2
Eurasia (CIS)	178.0	500	30	70	30	6.8	2.9	9.7
Developing countries	837.2[a]	250[b]	35	60	40	16.1	10.6	26.7
				TOTAL		32.7	15.9	48.6

Sources: WHO and authors' estimates.

l = liters; m³ = cubic meters

a. Based on a total population having access to safe water supply of 1,902.7 million people, with 44 percent of these receiving water through individual household connections.

b. This figure reflects a wide discrepancy among developing countries, from 100 l/capita/day for some utilities in the poorest countries or those experiencing severe water shortages to more than 400 l/capita/day in many megacities of Latin America and East Asia. The figure used in this calculation is a conservative average.

Source: Kingdom, Liemberger, Marin (2006)

In water and electricity delivery, what appears as an engineering problem really is a more complex phenomenon involving lack of maintenance resources and inadequate systems of incentives and accountability of management, staff and politicians (World Bank, 2004). Managers are risk averse, and prefer asking for more resources or closing financial gaps by cutting maintenance. Since there is no "ribbon-cutting" and pictures in newspapers, politicians are not interested in loss-reduction programs, and engineers may find more challenging designing new plants and expansion plans than fixing underground pipes. Reducing losses also generally involves unpopular activities such as cutting service to some users. Field staff also faces inadequate incentives in deciding their work effort when assigned to loss-reduction activities, as finding illegal connection may be dangerous for them, as well as maintaining inefficiency of this kind provides an opportunity to profit from corrupt practices. Hence, association with the private sector under the new performance-based service contracts is an alternative that has worked well in several places, such as Dublin, Sao Paulo, the state of Selangor (Malaysia) and Bangkok. (Kingdom, et al., 2006).

III. Microeconomic Aspects of Efficient Public Spending

Governments typically spend between 20 and 30 percent of GDP on goods and services. If they could increase the efficiency of the use of these resources, on a year by year basis, the impact on GDP would be substantial. The natural problem is defining and measuring efficiency, discussed in this section in two parts: a) allocative efficiency and the rate of return of public projects and programs; b) impact evaluation.

A. The rate of return of public projects

Besides doing things right (technical efficiency), the government has to allocate resources to the right things (allocative efficiency). If governments allocate their scarce resources to different activities based on the principle that the project's social rate of return exceeds the marginal social cost of public funds, and projects with the larger net social benefit are prioritized, then the project will have a positive impact on GDP. In practice, the comparison of benefits and costs must contemplate project specificity. This section discusses some of the factors that account for heterogeneity across projects, causing the threshold rate of return for efficiency to vary. It also shows some of the challenges that policymakers face when using rate of return estimates to guide their decisions.

1. Relationship with the marginal cost of public funds

Spending and taxing decisions are not independent because the marginal cost of public funds is a necessary element in the evaluation of public projects and programs. Since the decision rule is to undertake a project or program if the marginal social benefit exceeds the marginal social cost of public funds (Slemrod and Yitzhaki, 2001), then, estimating the marginal cost of funds is as important as estimating the benefit of projects. Public spending necessarily implies raising taxes, now or in the future. But taxation alters society's consumption and production decisions. Both the composition and levels of production and consumption bundles end up being different than they would be in the absence of taxation. Hence, society's welfare changes as a result of tax variations. The monetary value of the welfare change per unit of revenue raised is the marginal cost of public funds.

The shadow price of a dollar of public revenue is higher than 1 because, in addition to the dynamic efficiency costs described above, the government incurs administrative costs to collect taxes, which fluctuate between 1 and 4 percent of total tax collections (Gallagher, 2005: Warlters, et.al. 2005). Also, there are compliance costs that private agents incur to meet their lawful obligations, but there are no quantifications of this cost. Hence, each dollar of revenue collected by the government costs more than one dollar to society.

Most estimates of the marginal cost of funds are for developed nations, with relatively few calculations for developing countries. Estimates vary widely, according to the type of tax considered, the country, and the specific model used for the calculation. For instance, for the US, estimates range from 1.5 (Jorgenson et. al., 1990) to 2.7 (Feldstein, 1997). In developing countries, figures fluctuate from 1.5 for Cameroon to 2.2 in Bangladesh

(Devarajan et. al. 2001). For China, the sales tax deadweight loss estimate is 2.3 (Laffont et.al., 1997), while for India it ranges from 1.6 to 2.1 (Ahmad and Stern, 1987). Recent estimates for a large sample of African countries (Warlters et.al. 2005) indicate a range from 1.05 to 1.37. The details on estimating this cost are reviewed elsewhere (Devarajan, et. al., 2001) and for the discussion that follows, the important fact to bear in mind is that the marginal cost increases more than proportionately with the average tax rate. In fact, the welfare cost increases quadratically with the average tax rate (Stiglitz, 1988), which means that distortions or inefficiencies induced by the tax system increase more than the tax rate will.[18]

Given this range of estimates for the marginal cost of public funds, the natural question is what this implies for public expenditure evaluation and decision rules. Bevan (2007) presents a useful framework (summarized in Appendix 4) to analyze the required rate of return for various kinds of projects. To be efficient, a project's marginal social benefit must exceed the marginal cost of public funds, and the level of the required rate of return that meets the condition will depend on : 1) whether the specific project will require general tax revenue or will it be financed by user charges; 2) the extent to which the government will be able to appropriate some of the benefits of the project; 3) the values of some parameters such as the marginal cost of public funds, the recurrent expenditures generated by the project, and the rate of discount. For the average values of these variables considered originally by Bevan, the required ROR varies between 10 and 14 percent, judged to be a reasonable back-of-the-envelope estimation of a benchmark range.

This range spans a space of possible rates of return by considering variability in some of the main determinants. One of the elements that vary considerably with the type of project is the recurrent costs ratio (c). For instance, in a sample of World Bank projects, recurrent costs fluctuated from about 1 percent in energy projects to over 7 percent in education projects. In a sample of ADB projects, the highest recurrent cost ratios, of about 7 percent, were for health projects (Hood, et.al. 2002). More disaggregated studies report even higher recurrent cost ratios: roads fluctuate from 3% to 14%, general hospitals and urban health centers around 18 percent, and primary and secondary schools range from 6 percent to 70 percent of investment costs (Heller, 1982). Another source of variability is the rate of discount that projects might use. While health or environment projects might use low discount rates (in the 3 percent range), infrastructure projects might use discount rates towards the higher end of the spectrum. In addition, the MCF is different across countries, and as it rises, so will the rate of return required for economic efficiency. Additionally, the dimensionality of the problem increases if you consider that these variables might change through time. Table 7 shows different required rates of return that vary with the marginal cost of funds, the discount rate, and the recurrent cost ratio. These variables define the space of action of the public sector to have a positive

[18] The welfare cost of taxation, or the burden of taxation, will also depend on the relationship between the marginal tax rate and the average tax rate, which reflects the progressivity of the tax system. These complications are not treated here. Niskanen (2003) presents a short discussion with some back-of-the-envelope calculations for the US economy.

impact on GDP. This constellation of rates define the fiscal space of action for the government to have a positive impact on GDP through its spending.

Table 7
Required Rates of Return for Alternative Rates of Discount (r), Recurrent Cost Ratios (c), and Marginal Cost of Funds (MCF)

Project pays for itself (no taxes)

	r		
	0.03	0.05	0.09
c			
0.03	0.06	0.08	0.12
0.05	0.08	0.10	0.14
0.12	0.15	0.17	0.21

Project needs full tax financing

MCF = 0.15

	r		
	0.03	0.05	0.09
c			
0.03	0.07	0.09	0.14
0.05	0.09	0.12	0.16
0.12	0.17	0.20	0.24

MCF = 0.25

	r		
	0.03	0.05	0.09
c			
0.03	0.08	0.10	0.15
0.05	0.10	0.13	0.18
0.12	0.19	0.21	0.26

MCF = 0.4

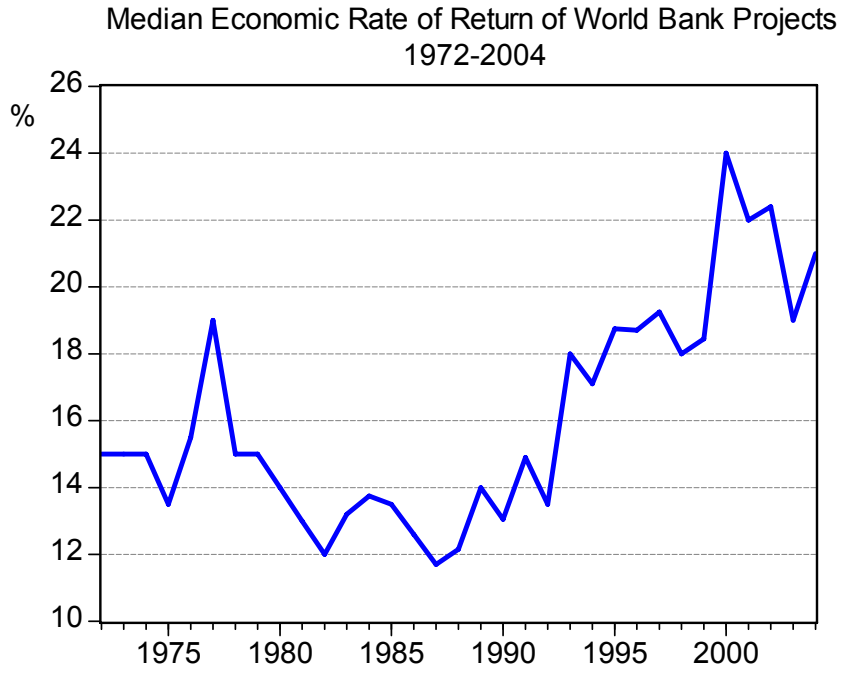
	r		
	0.03	0.05	0.09
c			
0.03	0.08	0.11	0.17
0.05	0.11	0.14	0.20
0.12	0.21	0.24	0.29

Source: calculations based on Appendix 4

2. The observed rate of return of public programs and projects

Obtaining historical information on the rate of return of public spending is a difficult task, but the data limitation may be partially overcome by considering a sample of public projects financed by the World Bank in recent decades. In spite of the limitations of this sample, derived from its small size in relation to public spending in each country and possible selection bias discussed below, there are interesting facts of the rate of return during the period 1980-2004. The observed rate of return of public spending changes through time and falls within the ranges estimated in the previous section. Naturally, this refers to the median of the sample for each year, and there is much variability across sectors and regions (Herrera, 2005). While the economic rate of return fluctuated between 12 and 14 % during the eighties, there is an upward trend during the nineties to settle around 18 or 20%.

Figure 13



Source: Calculation based on OED database

The upward trend in the rate of return of public spending is associated with a more stable macroeconomic environment. During the nineties, there was a generalized opening of trade and improvement of fiscal positions that are associated with better project performance. Different studies show that there is a statistically significant relationship between the macro environment and project performance (Isham and Kaufmann, 1999; Herrera 2005). This fact reinforces the point that fiscal policy does not act in a vacuum, and better project outcomes will be achieved in settings in which markets are left to operate so that prices are effective signals of fundamentals.

The median rate of return of about 20 percent seems high, despite the positive effect of improved macroeconomic fundamentals during the nineties. Though there might be a few public projects that have such rates, or even higher, it seems very unlikely that the typical public project yields this result (Harberger, 2005; Devarajan, et.al., 1997). If the typical public project or program produced these results, then almost any public spending would have a positive impact on GDP. Though estimates obtained from this database may reflect problems specific to it, probably the determinant ones, discussed below, may be applicable to the practice of cost-benefit analysis of public projects in other settings.

First, the portfolio has a high variability, with a standard deviation of around 20 percent. Hence, there are many bad public projects that the median-based analysis does not incorporate. The same program may have positive results in one location but negligible ones in a neighboring one. Program design and implementation are critical, and they are likely to change from one location to another, and through time. Finally, the median is unweighted, treating alike both small and big projects providing limited information on the overall return of public spending.

Second, there might be a selection bias if governments choose projects with the highest rate of return to seek funds from international organizations. Though the evidence for this hypothesis is not compelling, there are claims of different rates of return for projects funded by outside donors and those funded from the national budget (Devarajan, et.al. 1997). Other indirect evidence of the "cream-skimming" hypothesis would be the finding of rates of return on projects funded with EBRD resources systematically higher than those financed by the World Bank. In turn, projects funded by both institutions registered higher rates of return than those funded with EU cohesion funds (Florio, 1999).

And third, the data reveal an "optimism bias", as ex-ante (at appraisal) rates of return are, on average, higher than ex-post (at completion) rates. The bias is not constant, varying across regions and sectors, with projects in the urban development and infrastructure sectors showing the highest bias. The optimism bias is not exclusive of this database, as it has been detected some time ago (Harberger, 1997). The UK Treasury's guide to project evaluation, The Green Book, refers to this problem and proposes methods to incorporate the bias in project analysis, and the appraisal of EU Cohesion-funded projects seems to have similar problems (Florio, et.al. 2004). These results indicate that the profession needs to improve on the practice of evaluating the benefits of programs and projects, which leads to the following section.

B. Impact evaluation - the quest for more evidence-based economic analysis and policymaking

The previous section revealed the need of exploring deeper the cost and benefit estimation of traditional project evaluation. Economic analysis of projects would benefit from paying more attention to specifying the counterfactual, which would allow comparing the benefits from the program with those under an alternative scenario (Devarajan, et.al., 1997). This would reduce the optimism bias and lead to more credible evaluations and hence, better policies. The recent impact evaluation methods play a

fundamental role in this process. This section briefly presents the method, summarizes some results from case studies directly related to capital formation, and finally presents some caveats when drawing lessons for policy from these evaluations.

1. Method[19]

An impact evaluation assesses changes in an outcome indicator (education attainment, health, income level, or well-being) of individuals, households, communities, or firms, that can be attributed to a particular project, program or policy. The central impact evaluation question is what would have happened to those affected by the intervention if they had not in fact been affected by it. Since this is unobservable, the key challenge is to specify a counterfactual – that is, to select a group which is as similar as possible (in observable and unobservable dimensions) to those affected by the intervention. This comparison allows assessment of causality – attributing observed changes in the outcome indicator to the program, while removing confounding factors.

There are alternative methods to specify the counterfactual, that is, the selection of the comparison group that will be used to measure the impact of the intervention. Which method is chosen will depend on numerous considerations, such as the type of project, the evaluation question, the time constraint, data availability. The methods have different advantages and disadvantages that have been reviewed and presented in detail elsewhere (Baker, 2000 and Ravallion, 2001, 2005). Here, just an intuitive summary follows.

The core of impact evaluation is to select a comparison group, which should be identical (or as similar as possible) to the treatment group, with the key difference that the comparison group did not benefit from the program. The design of the comparison group is oriented to control for any bias due to differences in observable characteristics (schooling, geographic location, participation in the labor market, etc.) or unobservable ones (skill, motivation, connections, etc.) between the two groups. There are alternative methods to assign individuals into the treatment (beneficiary or affected by the intervention) and control (unaffected by the intervention) groups.

The first method, randomization, allocates the intervention randomly among eligible beneficiaries. The random selection of treatment and control groups ensures there is no bias, and hence, it is considered the most robust evaluation method. The disadvantages of the method emerge from the likelihood of its applicability due to the political circumstances that enable to select randomly beneficiaries for a program, or to ethical considerations that allow the policymaker to deny treatment to somebody that actually needs it. The method is also inapplicable in projects that have national coverage due to the impossibility of selecting a comparison (no-treatment group). This method is also known as an experimental design of the evaluation.

Other methods, known as non experimental, generate the comparison group that resembles the treatment group through econometric techniques. The selection into the

[19] This section draws on material presented in more detail in the impact evaluation website of the World Bank, on the Handbook on this topic (Baker, 2000), and several papers by M. Ravallion.

groups is generally done after the intervention. The most common approaches are the matching method, the double difference method, and the regression discontinuity method. The matching method attempts to match the comparison to the treatment group based on observable characteristics. The idea is to select as a comparison group those with the same probability of participation in the program. The limitation of this method arises from the assumption that there is no unobserved heterogeneity, that is, there is no bias arising from unobserved factors. The double difference method overcomes this limitation by comparing the differences between the treatment and comparison groups (first difference) before and after the program (second difference). The limitation of this method arises from the assumption that the bias arising from unobservable factors remains constant through time. Finally, the regression discontinuity method may be applied when a program is assigned based on a specific threshold which produces discontinuity in the participation (individuals participate or not) but the individuals just below the threshold are identical to those just above the threshold.

2. Public expenditure and capital formation- results from some impact evaluations

Most impact evaluations have centered on quantifying the impact of policies on human capital formation (education, nutrition and health) while relatively fewer studies have analyzed the impact of infrastructure or the impact of policies on building assets. This section describes some results of different evaluations that highlight the importance of public programs on capital formation.

There have been numerous evaluations performed on the policy impact on human capital formation, and currently the World Bank (2007b) is undertaking a project to extract lessons from several evaluations of a specific type of program[20]. Here the focus will be in the Mexican Progresa-Oportunidades, because it was the first program of its type, and the positive results of the evaluation (Skoufias, 2005) have led to its extrapolation to other countries (Brazil, Colombia, Bangladesh, Australia adopted this type of program for indigenous peoples, and New York City started one too) (World Bank, 2007b).

With variations, these programs transfer cash to poor families, conditioned on their undertaking certain actions to build up human capital of their children. Generally, these actions entail ensuring school attendance, visit to doctors and intaking a minimum amount of a nutritional supplement. The transfer is an incentive to build up human capital and, hence, is expected to alleviate long-run poverty. The program also has a short-run effect on poverty through the cash transfer income itself. However, the cash transfer may also affect negatively the beneficiary's labor supply on two counts: the income effect of the transfer will increase demand for consumption and leisure; and a substitution effect that may divert time towards meeting the program conditionality.

The Progresa impact evaluation shows positive results of the program (Skoufias, 2005). On the education side, the program led to an increase in school attendance for boys and

[20] The specific type of program is the conditional cash transfer, or CCT, described below.

girls in secondary and high school level (ages 12 to 17), with the effect being larger in girls: while boys attendance increased by 8 %, that of girls rose by 14%. There was no impact of the program in younger ages, probably due to the fact that school attendance for those ages was already high prior to the program. This led to a reallocation of resources of the education grant towards high school aged participants. In terms of health, Progresa reduced the incidence of illness for 0-2 year olds and 35 year-olds. In terms of nutrition, Progresa led to higher growth and reduction in the probability of stunting in children 12-36 mos. The evaluation indicated that labor market participation had not been affected by the program (Skoufias, et.al, 2007). This particular result may not be generalized, as impact evaluations of other programs show that the labor supply effect of transfers is significantly negative, as in the case of a rice subsidy in Sri-Lanka (Sahn et.al. 1996). It would be interesting to compare both cases to understand the different response of labor supply to the transfer.

A major puzzle remaining from this evaluation is the lack of impact of the program on learning scores (Skoufias, 2005). If the quality of education is the critical component in human capital, as argued in previous sections, it is legitimate to question the impact of the program even if more children went to school. There are not many evaluations of impact of policies on learning scores, but there are a few that are worth highlighting. One of the better known cases is the Colombia voucher program (an in-kind transfer and not a CCT), probably due to the method used (randomization) and because the cognitive tests were administered 4 years after the initiation, considered a long-enough lag to observe learning results (Angrist, et. al. 2002). The problem with the evaluation was that the sample was very small (aimed only at a specific cohort of students that applied in a specific year) and that the test was specifically designed for the evaluation and hence might have been perceived as an academic exercise. Results showed an increase in learning scores of .2 standard deviations of program participants. In a posterior evaluation, researchers were able to match program applicants with their college entrance examinations (seven years after their application to the program). All the previous results were verified with a much longer lag and a larger sample (Angrist, et. al.2006). Another impact evaluation shows a positive response of learning scores to changes in the teacher remuneration system that reduced absenteeism in India (Duflo et.al 2006).

Besides the impact on human capital formation, some evaluations have also analyzed the impact of transfers on family savings and investment in productive assets. Another evaluation of the Mexican Progresa (Gertler. et. al , 2006) showed that five years after the program, rural household beneficiaries had invested in farms and in microenterprise activities. Increased investment in productive activities generated a permanent increase in income that allowed an increase in consumption of 34%.

This reaction of beneficiaries is not generalized; rather, the measured effects of public spending depend upon program design, on individual's preferences, in particular their risk-aversion, and on the timing of the evaluation. A contrasting example to the Mexican case is a public program in China evaluated 5 years after program initiation (Ravallion and Chen, 2005), and then 5 years after the program completion (Chen, Mu and Ravallion, 2006). The program, implemented in 1,800 villages on south-west China, had

two components: one that was expected to affect equally all households, consisted in improving infrastructure such as rural roads, power lines and piped water. The other component was a loan to households to finance initiatives to raise farm yields, animal husbandry, and tree planting. The medium term evaluation, based on household surveys five years after the program initiation, showed that incomes had risen but consumption did not change. The long term evaluation, based on survey data of 5 years after program completion (10 years after program initiation), showed that the permanent income of the households had not changed and neither had consumption. Whether these results can be interpreted as evidence of the permanent income hypothesis of consumption, as the authors did, or evidence of extreme risk aversion of the poorest households, is debatable. The main lesson is that program design will affect the outcomes and must consider the beneficiaries' preferences and constraints to anticipate their reaction to the public program.

The role of public spending on physical capital formation, and its consequences for welfare and poverty, has not been explored extensively using impact evaluation methods. Curiously, the cost-benefit analysis of these projects registers the highest difference between ex-ante and ex-post rates of return in projects funded by the World Bank (Herrera, 2005). Similarly, the UK Treasury pinpointed capital expenditures as the largest source of optimism bias in public investment projects, and identified as its causes the poor definition of the objectives of the project and the poor identification of stakeholder requirements (UK Treasury, 2003). On both counts, identifying the projects objectives and identifying stakeholders, the impact evaluation methods can shed much light.

Only until recently have some infrastructure projects involved impact evaluations. Some examples are in the area of water and sanitation (Poulos, et.al, 2006.), in urban transport (Boarnet, 2006), in rural roads (van de Walle, 2006), and in dams (Duflo, 2005). These evaluations have in common that they focus on two key issues: endogeneity of placement and the distributional consequences. For instance, the urban transport project evaluation incorporates the fact that city dwellers adjust their location and travel patterns partially based on public transport considerations. More motivated and employable people might choose to live near the transport route, and hence ex-post evaluations that find that people living next to the project have higher employment rates must take that into consideration. In the case of dams, the evaluation incorporated the fact that these are placed depending on the regions' wealth and was able to identify the counterfactual group with geographic characteristics. The evaluation showed the different impact of dam construction in the own district and the downstream district: agricultural output increased downstream and poverty fell, while in the districts where the dam was built poverty increased. The evaluation showed a net present value of 1 percent, without considering the deadweight cost of public funds required for project funding.

3. Lessons from impact evaluations

All evaluation methods have advantages and disadvantages. While there are advocates of specific evaluation methods (i.e. Duflo, 2004), others propose a flexible approach so that the evaluation can be adapted to the problem, to the setting, and to the availability of data (Ravallion, 2005). Regardless of the position in the debate, the underlying assumption is that the average treatment effect is something useful for the policymaker when considering scaling up or extrapolating results from evaluations.

When attempting to extrapolate results from a specific evaluation, the policymaker must be aware of some common problems to most evaluations. The first one, the publication bias, refers to the preference of researchers and editors for statistically significant results. The bias has been well documented in empirical research in economics, and other fields such as medicine and social science (De Long, 1992; Stanley, 2005). Hence, published results must be interpreted with caution as the impact of the intervention will tend to be overestimated. In the particular area of evaluation there is not much evidence of the bias. However, a study that compared different evaluation methods found that papers reporting "no-impact" results were always easier to replicate with all the methods, while the positive impact results were difficult to replicate. Additionally, papers using popular evaluation methods, the non experimental type, also registered a higher impact than studies using the experimental approach, by a margin that could be as high as 10% of the income of the target group (Glazerman, et.al. 2003).[21] The potential for this bias has led some researchers to advocate for institutions to ensure that the negative results are systematically disseminated (Duflo, 2004).

A second problem that may affect evaluations of the impact of policies is that the agents being tested (treatment group), or the control group, may change their behavior transitorily due to the assessment being performed. Due to the increased attention or the fact that beneficiaries are receiving something during the assessment, individuals may receive a temporary boost to their motivation, which will bias results toward s finding a positive impact of the intervention (this is called the Hawthrone effect). The control group may also react to not having been included in the program and may boost performance to compensate for the fact (this is the John Henry effect). It is important that the evaluation take into consideration these effects. For instance, instead of having two groups (treatment and control) a third one can be added that receives some benefit, though not as effective as the experimental intervention being tested.

A third factor to consider when extrapolating results from an evaluation, especially for scaling up a program, is its general equilibrium implication. It is essential to consider the individuals' incentives and reactions to the program. For instance, if a small scale school voucher program (like the Colombia case) were to be scaled up, that would affect the functioning of the whole school system in ways that will completely alter results. For instance, if schools are provided with incentives to improve outcomes, and they are allowed to choose their students, they will react by being more selective ("cream-

[21] These results refer only to impact evaluations of programs to increase earnings incomes of target groups. However, the Galzerman et.al. paper is the only published paper that has done this analysis.

skimming") which will lead to different test scores between private and public schools. Schools will also tend to spend more money in what parents value more, such as having good looking infrastructure, but that not necessarily improves academic achievement. The Chilean case provides evidence of how voucher programs led to higher school segregation with little impact on learning scores but higher parent satisfaction due to the increased choice (Hsieh and Urquiola, 2003).

The scaling up of programs to the national level may affect relative prices and may require additional taxation. Both effects mitigate the impact of the original policy, leading to an overestimation of benefits based on partial equilibrium analysis. For example, a massive educational program will affect educational skill composition of the labor force, which in turn will lower skill premiums and affect the wage structure. When the general equilibrium considerations of a tuition subsidy program include the effect of the additional taxation required to finance the scaling up, the overestimation of the effects may reach up to 10 times the general equilibrium impact (Heckman et al., 1998). Other comparisons of benefits estimated in a general equilibrium setting versus a partial equilibrium analysis, indicate that the overestimation of the impact due to ignoring the relative price changes and the subsequent household adjustment to them is extraordinary, ranging from 4 to 10 times the more realistic impact (Sieg, et.al., 2000).[22] The necessity of considering the general equilibrium highlights the complementarity of detailed micro evaluations with more structural economic modeling (Bourguignon, 2004).

The fourth element to be considered when extrapolating results is that the institutional context determines the magnitude of the impact. One program can have impacts in one location but not in another. The Food for Education assessment in Bangladesh (Galasso and Ravallion, 2005) showed that the program's results were correlated with observable village characteristics. These observable characteristics reflected features of the local decision-making process. For instance, the anti-poverty program had less impact in villages with greater inequality, or that were more isolated. Inequality led to targeting of expenditure towards different groups that would not benefit as much, while isolation probably reduced the perceived accountability of local decision-makers.

IV. Political Economy of Government Expenditure

This very brief section acknowledges that institutional factors, both related to electoral institutions and to budgetary processes, will affect the size, predictability, composition,

[22] The Sieg et. al. paper does not refer explicitly to public spending, but to a policy aimed at achieving a cleaner environment. The partial equilibrium estimation of the benefit involved comparing willingness to pay (WTP) measures in a community before and after the policy was in place. However, this method ignores that people can choose to move from one community to another, in response to the cleaner air. People will tend to relocate to the communities with cleaner air, leading to rising housing prices in those communities, and vice versa in communities with lower air quality. Additionally, the households will not be the same (they have different preferences) in a community in the two distinct points in time (before and after). Considering the differences in preferences and incomes that the different household compositions imply, the willingness to pay measures that incorporate changes in housing prices reveal much lower benefits.

and the extent of evaluation and monitoring of public spending. These were the topics discussed in previous sections from a technical standpoint and it would be naïve to conclude the paper without recognizing the role of political economy in determining these outcomes.

Rather than thinking in terms of controlling political indiscipline, it may be more constructive to consider the types of institutions that will improve the quality of decisions, recognizing that politicians have different incentives and horizons than technocrats. Politicians have allegiances with specific geographic constituencies, but also with national social groups (Milessi-Ferreti, et.al.2003). And how will they be held accountable by both depends on the political competition and accountability rules that the electoral system defines.

Electoral rules will affect public spending to the extent that they affect political competition and accountability. Under the plurality rule system (majoritarian rule) the candidate with the largest number of votes wins a seat in parliament. This system maximizes individual political accountability, and hence tends to reduce the size of spending and waste or the ability of politicians to extract rents. However, it will reward politicians for channeling funds to specific regions and will affect the composition of spending in favor of goods or services that are more easily targeted towards specific groups at the expense of broad or more universal spending programs. The reverse will happen in proportional representation systems, in which candidates are drawn from national party lists and are elected based on shares of total voting. This system reduces individual accountability and evidence shows a positive association with higher corruption and waste (Persson, Tablellini and Trebbi, 2000). Evidence also shows that proportional systems tend to spend more on transfers or the universal good, while majority rule systems spend more on goods or services that are easier to target geographically (Milesi-Ferretti, et.al, 2003).

At the center of this discussion lies the common pool problem in public budgeting: general tax funds are used to finance projects that benefit particular groups of individuals that perceive the full benefit of a project but only a fraction of the cost. Hence, institutions that facilitate the internalization of the entire budget constraint by politicians and decision makers would lead to better fiscal outcomes. In this sense, the fragmentation of the budget has to be minimized, and fragmentation occurs in different instances: 1) when there are off-budget items; 2) when there is indexation of spending or entitlements with parameters fixed by law; 3) when there are mandatory spending laws or revenue earmarking laws; 4) when the use of contingent liabilities and public guarantees affect future budget allocations and take precedence over future decisions (Von Hagen 2002).

Because spending predictability and efficiency require sustainable policies, it is crucial to balance between delegation and representation. Delegation refers to the power or authorization to an agent to act as a "fiscal entrepreneur". Evidence shows that centralizing expenditure decisions in the ministry of finance is associated with positive fiscal outcomes (Alesina, et.al. 1999) Similarly, the "growth champions" identified in

the first section of this paper have in common centralized fiscal systems. However, adequate representation is required for sustainability.

Given the high welfare cost associated with fiscal policy volatility, it would be beneficial to consider what kind of institutions would make policy more predictable. Both fiscal and electoral institutions have an impact on fiscal outcomes and on public spending, and it would be useful to examine why the EU countries have lower expenditure volatility. While some researchers focus on the role of electoral institutions (Person et al. 2007 and 2005), others concentrate on the budgetary process and fiscal institutions (Von Hagen, 2002). Though there is in general a positive assessment of fiscal rules, this is still a matter of debate. For instance, adjusting fiscal variables for the business cycle to focus on the structural fiscal balance or the discretionary component of fiscal policy seems impeccable from the theoretical standpoint. But, in practice, ex-post evaluation of the cyclical adjustments in the EU countries shows that the output gap is systematically underestimated, leading to more expansionary fiscal policy (Ley, 2006; Vidal, 2007). If the economists' toolkit leads to these systematic mistakes in developed nations, it is legitimate to ask about the validity of its application in developing countries. Some propose complementing existing rules by setting expenditure growth ceilings, but this too is subject to debate (Von Hagen, 2005). What is needed is national consensus on the importance of reducing fiscal volatility. This is also crucial for ensuring efficiency of public spending.

V. Conclusions

Public spending has the potential of affecting capital formation and GDP. If the specific project meets the economic efficiency criterion, it can help remove "growth bottlenecks" or mitigate consequences of market failures. But public spending is also subject to failure and society incurs costs because of spending. These derive mainly from the distortions arising from taxation or from the volatility induced to household consumption.

Expenditure decisions are not independent from taxation. From the micro standpoint, taxation levels determine the marginal cost of public funds, which is the minimum threshold for a project's rate of return for it to satisfy economic efficiency criteria. If public resources are allocated to projects that satisfy this requirement, and spending is prioritized based on the projects' rate of return, then public spending will have a positive impact on GDP.

From the macro viewpoint, prolonged spending expansions beyond the stable funding source of taxation expose expenditure to the volatility of financial markets. Not surprisingly, public spending is more volatile in highly indebted less developed economies. The stop-go cycles of expenditure induce volatility to household consumption that imposes welfare costs for society, estimated at the equivalent of 8 percent of consumption per year in developing economies. The burden of this cost is most likely borne by the poor, which have subsistence consumption levels and lower tolerance to risk.

If the expected benefits must meet such hurdles, then gauging the benefits deserves special attention. At the macro level, inferring the productivity of public spending or its impact on capital formation requires considering jointly the impact of additional spending with the cost of taxation, differentiating between spending and capital formation, and recognizing heterogeneity in the quality and productivity of capital. These considerations lead to more realistic assessments of the impact of public spending on growth.

This paper highlights the need of going beyond the prescribed country specificity. It is essential to consider different types of capital, and even going even further, it is essential to go to project level specificity. Project-specific information, together with macroeconomic variables such as the marginal cost of public funds and the rate of discount, will determine a constellation of required rates of return that government projects must satisfy to have a positive impact on GDP. This is the fiscal space.

At this level, performing credible and rigorous project evaluations is critical. In this respect, impact evaluation is a complementary tool that can provide empirical support to quantifying the benefits, making evaluations more robust. From individual impact evaluations, public expenditure analysis has to return to the general equilibrium setting, to rank and choose among the alternative uses of public funds. Only with an operational project evaluation system will the efficiency of public spending be enhanced.

In deciding expenditure levels, composition, and efficiency enforcement, politicians will be involved. Hence, institutions that will improve decision making with regards to spending must consider that incentives and horizons of politicians and technocrats are different. Institutions that reduce fragmentation of the budget process will lead to better outcomes, and these will be sustainable with adequate political representation. Predictability and transparency of spending, with increased accountability of both technocrats and politicians, will lead to better outcomes of public spending.

References

Adam, C. and D. Bevan, 2005. "Fiscal deficits and growth in developing countries", *Journal of Public Economics*, 89 pp 571-597

Afonso, A. and M. St Aubyn 2004 Non-parametric approaches to education and health: Expenditure efficiency in OECD countries". Mimeo. Technical University of Lisbon.

Afonso, A. L. Skuchnecht, and V. Tanzi , 2006. Public sector efficiency: evidence for new EU member states and emerging markets. European Central Bank. Working Paper Series. No. 581. January

Agenor, P. and K. Neandis, 2006. "Health, education and growth with public infrasturcture." Mimeo.

Agenor, P., E. Pinto-Moreira, and N. Bayraktar, 2007. "Allocation of public expenditure and growth: operational framework for the Haiti PEMFAR report". The World Bank. Mimeo. March.

Ahmad, E. and N. Stern, 1997. "Alternative sources of government revenue: illustrations from India, 1979-80", in D. Newberry and N. Stern eds. *The Theory of Taxation for Developing Countries*, Oxford University Press, New York.

Alesina, A., R. Hausmann, R.Hommes and E.Stein, 1999 "Budget institutions and fiscal performance in Latin America," *Journal of Development Economics*, August 59: 233-53.

Alesina, A. and G. Tableeini, 2005. "Why is fiscal policy often procyclical?" NBER Working Paper 11600 September.

Angrist, J. E. Bettinger, E. Bloom, E. King, and M. Kremer 2002 "Vouchers for private schooling in colombia: eveidence from a randomized natural experiment", *American Economic Review* 92 No. 5 December 1535-1558

Angrist, J. E. Bettinger, and M. Kremer (2006) "Long-term educational consequences of secondary school vouchers: evidence from administrative records in Colombia", *American Economic Review* vol. 96 No. 3 June 847-862

Arestoff, F. and C. Hurlin (2006) "Estimates of government net capital stocks for 26 developing countries, 1970-2002", The World Bank Working Paper Series WPS 3858, March.

Atkinson, A., S. Burgess, B. Croxson, P. Gregg, C. Propper, H. Slater, D. Wilson (2004) Evaluating the impact of performance-related pay for teachers in England. CMPO Working Paper 04/113. University of Bristol. December.

Baker, J. 2000 *Evaluating the Impact of Development Projects on Poverty A Handbook for Practitioners*. The World Bank.

Bevan, D. 2007 "Fiscal policy for growth with stability", Mimeo. Background paper for World Bank project on Fiscal Policy for Growth and Development.

Bleaney, M. N. Gemmell and R. Kneller 2001 "Testing the endogenous growth model: public expenditure, taxation and growth over the long-run", *Canadian Journal of Economics*, 34, 1, pp 36-57.

Boarnet, M. "Conducting impact evaluations in urban transport". The World Bank Doing Impact Evaluation Series No. 5

Bourguignon, F. ,2004. "Comments on Attanasio and Duflo" in *Accelerating Development*, ed. F. Bourguignon and B. Pleskovic World Bank and Oxford university Press.

Bourguignon, F., C. Diaz-Bonilla, and H. Lofgren 2006 "Aid, service delivery and the MDG in an economy-wide framework", Mimeo. December.

Calderon, C. and L. Serven, 2003. "The output cost of Latin America's infrastructure gap", *The Limits of Stabilization; Infrastructure, Public Deficits, and Growth in Latin America*, ed. W. Easterly and L. Serven. Stanford University Press and The World Bank.

Calderon, C. and L. Serven, 2004. "The effects of infrastructure development on growth and income distribution" The World Bank Policy Research Working Paer Series WPS 3400 September.

Canning, D. and E. Benathan, 2000. "The social rate of return on infrastructure investments", The World Bank Research Paper Series WPS 2390 July.

Chaudhury, N., J. Hammer, M. Kremer, K. Muralidharan and H. Rogers (2006) "Missing in action: teacher and health worker absence in developing countries", *Journal of Economic Perspectives*. Vol.20, No. 1, Winter pp.91-116.

Chen, S. R. Mu and M. Ravallion 2006 "Are there lasting impacts of aid to poor areas? Evidence from rural China". The World Bank Working Paper Series WPS 4084. December.

Chong, A. and M. Gradstein 2006. "Policy volatility and growth" IADB Research Working Paper Series 578. August.

Colletaz, G. and C. Hurlin (2006) "Threshold effects of the public capital productivity : an international panel smooth transition approach" Laboratoire d"Economie d'Orleans Document de Recherche No. 2006-01.

Debreu, G. (1951) "The coefficient of resource utilization", *Econometrica*, 19 pp. 273-292

Devarajan, S., L. Squire, and S. Suthiwart-Narueput, 1997. "Beyond the rate of return: reorienting project apparaisal", *The World Bank Research Observer*, vol. 12, No. 1 February.

Devarajan, S.,S. Suthiwart-Narueput, and K. Thierfelder 2001 "The marginal cost of public funds in developing countries", *Policy Evaluations with Computable General Equilibrium Models*, Routledge Press, London and New York.

Duflo, E., 2005. "Dams", NBER Working Paper 11711 October.

Farrell, M. 1957 "The measurement of productive efficiency", *Journal of the Royal Statistical Society, Series A*. Vol. 120, No. 3, pp. 253-290.

Fatas, A. and I. Mihov, 2003. "The case for restricting fiscal policy discretion", *Quarterly Journal of Economics*, November.
Fatas, A. and I. Mihov, 2004 "Fiscal discipline, volatility and growth". Mimeo. INSEAD and CEPR.
Fatas, A. and I. Mihov 2005. "Policy volatility, institutions and economic growth". CEPR Discussion Paper 5388.

Feldstein, M. 1997. "How big should government be?", National Tax Journal, Vol. 50, No.2, June pp. 197-213

Fernald, J., 1999. "Roads to prosperity? Assessing the link between public capital and productivity" *American Economic Review,* vol. 89, No. 3 June pp. 619- 638

Filmer, D., J. Hammer and L. Prtitchett 2000 "Weak links in the chain: a diagnosis of health policy in porr countries", *The World Bank Research Observer* vol.15 No. 2 August

Florio, M. and S. Vignetti 2004. "Cost-Benefit analysis, development planning and the EU Cohesion Fund : learning from experience" Working paper 31.20004 Di[partamento di Economia Politica e Aziendale, Universita degli Studi di Milano. November.

Galasso, E. and M. Ravallion, 2005. "Decentralized targeting of an antipoverty program", *Journal of Public Economics*, 89, pp. 705-727.

Gali, J. , 1994. "Government size and macroeconomic volatility", *European Economic Review*, 38, pp. 117-132

Gallagher, M. 2005 "Benchmarking tax systems", *Public Administration and Development*, 25 pp. 125-144

Gertler, P., S. Martinez and M.Rubio-Codina 2006. "Investing cash transfers to raise long-term living standards". The World Bank Working Paper WPS 3994. August.

Glazerman, S.,D. Levy, and D. Myers 2003. "Nonexperimental versus experimental estimates of earnings impacts", *The Annals of the American Academy of Political and Social Science* 589 September

Glomm, G. and F. Rioja, 2004 "Fiscal policy and long run growth in Brazil". Mimeo. World Bank.

Gramlich, E., 1994. "Infrastructure investment: a review essay", *Journal of Economic Literature,* vol 32 no. 3 pp. 1176-1196.

Gupta, Sanjeev, and M. Verhoeven. 2001. "The efficiency of government expenditure: experiences from Africa". *Journal of Policy Modeling*, 23, 433-467.

Hanushek, E. and D. Kimko (2000) "Schooling, labor force quality and the growth of nations", *American Economic Review* 90 No. 5 Becember pp. 1184-1208

Hanushek, E. and L. WoBmann (2007) The role of education quality in economic growth. The World Bank. Working Paper Series WPS 4122. February

Harberger, A. 1997 "Economic project evaluation, part I : some lessons for the 1990s", *Canadian Journal of Program Evaluation*. Special issue pp. 5-46

Harberger, A. 2004 Project evaluation as development priority. Mimeo. The World Bank Conference on Public Expenditure Efficiency and Growth. August.

Harberger, A. 2005. "On the process of growth and economic policy in developing countries" PPC Issue Paper No. 13. USAID Development Experience Clearinghouse. December

Heckman, J. L. Lochner and C. Taber 1998. "General-Equilibrium treatment effects: a study of tuition policy", *American Economic Review*, Vol. 88, No. 2 pp. 381-386. May
.
Herrera, S. 2005 "The economic rate of return of World Bank projects" The World Bank. PRMED. Mimeo. September.

Herrera, S. and G. Pang 2006 "How efficient is public spending on education?", La Educacion en Colombia, Banco de la Republica

Hnatkovska, V. and N. Loayza (2003) Volatility and growth. Mimeo. The World Bank. August.
Heller, P. 1982 "Recurrent costs. Problems in less developed counties", USAID Policy Paper . May
Hsieh, C. and M. Urquiola 2003. "When schools compete, how do they compete? An assessment of Chile's nationwide school voucher program" NBER Working Paper 10008. September.

Hood, R. D. Husband, and F. Yu 2002 "Recurrent expenditure requirements of capital projects", The World Bank Policy Research Paper WPS 2938 December

Hurlin, C.(2006) Network effects of the productivity of infrastructure in developing countries. The World Bank Working Paper Series WPS 3808, January

International Energy Association (IEA) 2002, 2006 *World Energy Outlook*
.
International Monetary Fund (IMF) 2005 "Public investment and fiscal policy-lessons from the pilot country studies", FAD. April.

Isham and Kaufman "The forgotten rationale for policy reform; the productivity of investment projects*", Quarterly Journal of Economics*, 149, 1.

Jacob. B. and S. Levitt, 2003. "Rotten apples: an investigation of the prevalence and predictors of techer cheating" *Quarterly Journal of Economic,* vol. 118, No. 3

Jalan, j. and M. Ravallion, 2001. "Behavioral responses to risk in rural China", *Journal of Development Economics,* vol. 66 pp.23-49

Jorgenson, D. K. Yun 1990 "Tax reforms and US economic growth", *Journal of Political Economy*, 88 pp 151-193

Kingdom, W.,R. Liemberger and P. Marin, 2006. "The challenge of reducing non-revenue water (NRW) in developing countries", Water Supply and Sanitation Sector Board Discussion Paper Series No. 8, December.

Kneller, R. M. Bleaney andN. Gemmell 1999 "Fiscal policy and growth: evidence from OECD countries", *Journal of Public Economics* pp. 171-190

Laffont, J.J. and C. Senik-Leygonie 1997 *Price Controls and the Economics of Insitutions in China* OECD Development Publications.

Ley, E. 2005 "Cual ha de ser el ciclo de referencia para los objetivos fiscales-el economico o politico?" Presentation at CEPAL conference. Santiago de Chile.

Laoyza, N., R. Ranciere, L. Serven, J. Ventura, 2007. "Macroeconomc volatilty and welfare in developing countries", Mimeo. World Bank. April.

Milessi-Ferretti, G.M., R. Perotti, and M. Rostagno 2002 "Electoral systems and public spending", *Quarterly Journal of Economics*. May.

Morduch, J., 1995. "Income smoothing and consumption smoothing", *Journal of EconomicPerspectives, Vol. 9, No. 3 Summer*

Niskanen, W., 2003. "The economic burden of taxation", paper presented in the FRB of Dallas conference *The Legacy of Milton and Rose Friedman's Free to Choose: Economic Liberalism at the Turn of the 21st Century*. October.

Ogaki, M. and Q. Zhang, 2001. "Decreasing relative risk aversion and tests of risk sharing", *Econometrica,* vol. 69, No. 2. March.

Patrinos, H., C. Ridao-Cano, c. Sakellariou (2006) Estimating the returns to education: accounting for heterogeneity in ability. The World Bank Working Paper Series WPS 4040, October

Persson, T., G. Roland, and G. Tabellini 2007 "Electoral rules and government spending in parliamentary democracies". Mimeo. IIES Stockholm University. January.

Poulus, C. S. Pattanayak, and K. Jones. 2006. A guide to water and sanitation sector impact evaluations. The World Bank Doing Impact Evaluation Series No. 4. December.

Pritchett, L. 1996 "Mind your P's and Q's. The cost of public investment is not the value of public capital", The World Bank Policy Research Paper WPS 1660

Pritchett, L. "The tyranny of concepts: CUDIE (Cumulated, Depreciated, Investment Effort) is Not Capital*", Journal of Economic Growth*, vol. 5 December pp 361-384

Pritchett, L. (2001) "Where has all the education gone?" *The World Bank Economic Review* vol 15, No. 3 pp 367-391

Psacharopolus, G. and H. Patrinos (2004) "Returns to investment in education: a further update" *Education Economics* vol 12, No. 2 pp.111-134.

Ravallion, M. 2001 "The mystery of the vanishing benefits: an introduction to impact evaluation", *The World Bank Economic Review*, vol. 15 No. 1 pp.115-140
Ravallion, M. 2005. "Evaluating anti-poverty programs", *Handbook of Development Economics* Volume 4 ed. T.P. Schultz and J. Strauss. Amsterdam, North Holland.
Ravallion, M. and S. Chen 2005. "Hidden impact? Household saving in response to a poor-area development project", *Journal of Public Economics* Vol. 89 pp. 2183-2204

Reinikka,R. and J. Svensson, 2001. "Explaining the leakage of public funds" The World Bank Policy Research Working Paper WPS 2709 November.

Sahn,D. and H. Alderman, 1996. "The effect of food subsidies on labor supply in Sri Lanka", *Economic Development and Cultural Change*, Vol. 45, No. 1, pp. 125-145, October

Sieg, H., V. Kerry, H. Spencer, and R. Walsh 2000. "Estimating the general equilibrium benefits of large policy changes : the clean air act revisited", NBER Working Paper 7744. June.

Skoufias, E. 2005 "PROGRESA and its impacts on the welfare of the rural households in Mexico", International Food Policy Research Institute Research report 139, Washington
Skoufias, E. V. di Marco 2006 "Conditional cash transfers, adult work incentives and poverty. Mimeo. LACEA Meeting. Mexico City.

Skoufias, E. 2007 The impact evaluation of the PRGRESA/Oportunidaedes program of Mexico. The World Bank. Mimeo. PRMPR. April.

Slemrod, J. and S.Yitzhaki 2001 "Integrating expenditure and tax decisions: the marginal cost of funds and the marginal benefit of projects", *National tax Journal* Vol. 54, No. 2 pp 189-202.

Stanley, T.D. "Beyond publication bias", *Journal of Economic Surveys* Vol. 19, No. 3

Stiglitz, J., 1988. *Economics of the Public Sector*. Second Edition W.W. Norton & Company.

Suescun, R. 2007. "The role of fiscal policy in human development and growth", LAC Regional Study. March
Talvi, E. and C. Vegh, 2005 "Tax base variability and procyclicality of fiscal policy", *Journal of Development Economics,* 78, pp 156-190.

Townsend, R., 1994. "Risk and insurance in village India", *Econometrica,* 62, pp 539-591.

Van de Walle, D. 2006. "Impact evaluation of rural roads". Mimeo DECRG. The World Bank. October

United Kingdom Treasury *The Green Book* available online http://greenbook.treasury.gov.uk/

Vegas, E. and L. Umansky (2005) *Improving teaching and learning through effective incentives: what can we learn from education reform in Latin America?* Washington D.C.

Vidal, J.P.2007 "What do economists really know about fiscal policy and economic fluctuations? The unpleasant journey from theory to practice" Presentation in Workshop on Fiscal Policy. Banca d"Italia, Perugia. March.

Von Hagen, J.2002 "Fiscal rules, fiscal institutions, and fiscal performance", *The Economic and Social Review* vol. 33 No.3 Winter pp. 263-284

Warlters, M. and E. Auriol 2005 "The marginal cost of public funds in Africa", The World Bank Research Paper Series WPS 3679 August.

WoBmann, L. (2003) "Specifying human capital", *Journal of Economic Surveys*. Vol 17, No. 3, pp.239-270

World Bank, 2002. *Peru: Restoring fiscal discipline for poverty reduction: Public expenditure review*. Report 24286-PE.
World Bank, 2004. *WDR 2004: Making Services Work for the Poor People* The World Bank and Oxford University Press.
World Bank. 2007a "Fiscal policy for growth and development- further analysis and lessons from country case studies". Mimeo. April.
World Bank. 2007 b "Fiscal policy and economic growth: lessons for transition economies", Europe and Central Asia Region, PREM, Mimeo April.
World Bank. 2007c "Tackling poverty in the short and long-run: an assessment of the experience of conditional cash transfer programs"

Appendix 1
The Cost of Output and Consumption Volatility in Developing Economies

Several papers have shown that output growth in developing economies is more volatile than in industrial economies (Agenor, et.al. 2000, De Ferranti, et.al. 2000). Fewer papers have shown that this volatility contrast is even stronger in household consumption, and furthermore, that consumption volatility is larger than output volatility in developing economies. Only recently has research incorporating this feature in economic modeling (Neumeyer et.al. 2005, and Garcia, et. al., 2006).

A natural question that emerges regards the welfare cost of volatility for households in developing countries. The answer will basically depend on the amount of risk and the degree of risk aversion of households. Since Lucas' quantification of this cost for the US (1987, 2003), there has been considerable development on this topic with limited application for developing countries. The object of this note is to present some estimates using a variant of a simple approximation suggested by Athanasoulis and van Wincoop (2000). The note is divided into three sections. The first one motivates the discussion by presenting data on volatility of output and consumption across the world. The second one presents rough estimates of the cost of output and consumption volatility. The third section concludes and proposes a work program for further refinement of these estimates.

I. Volatility of Output and Consumption across the World

Table 1 shows volatility across different regions of the world. GDP is more volatile in ECA, MENA and SSA. At the regional level, only in LAC does the volatility of consumption exceed that of output (col. 7, Table 1). When country-level data is examined (in the next section) this does not hold, generally.

Table 1
Output and Consumption across regions of the World, 1961- 2005
(per capita growth rates)

	GDP			Consumption			
	(1)	(2)	(3)=(2)/(1)	(4)	(5)	(6)=(1)/(2)	(7)=((5)/(2)
	$\bar{x}$	σ	$\sigma / \bar{x}$	$\bar{x}$	σ	$\sigma / \bar{x}$	σ con/ σ gdp
East Asia & Pacific	5.24	4.11	0.79	4.26	2.81	0.66	0.68
Europe & Central Asia	0.31	5.01	16.24	0.81	3.80	4.71	0.76
European Monetary Union	2.71	2.15	0.79	2.72	2.08	0.77	0.97
Latin America & Caribbean	1.56	2.35	1.51	1.67	2.53	1.52	1.08
Middle East & North Africa	0.89	3.12	3.52	1.05	2.22	2.11	0.71
South Asia	2.56	2.56	1.00	1.46	2.44	1.67	0.95
Sub-Saharan Africa	0.61	2.11	3.45	0.53	1.69	3.21	0.80
World	1.92	1.38	0.72	1.81	1.12	0.62	0.81

By level of income, the HIPC countries show more aggregate volatility (Table 2). The curious result is that the volatility of consumption is greater than that of output when counties are aggregated by their income level, implying that there might be risk-pooling within regions but not within the income group.

Table 2
Output and Consumption across the World, by level of income 1961- 2005
(per capita growth rates)

	GDP			Consumption			
	(1)	(2)	(3)=(2)/(1)	(4)	(5)	(6)=(5)/(4)	(7)=(5)/(2)
	$\bar{x}$	σ	$\sigma / \bar{x}$	$\bar{x}$	σ	$\sigma / \bar{x}$	σ con/ σ gdp
Heavily indebted poor countries (HIPC)	0.13	1.80	13.57	-0.25	2.06	-8.09	1.15
High income	2.61	1.58	0.61	2.54	1.36	0.54	0.86
Least developed countries: UN classification	0.44	1.93	4.43	-0.46	2.70	-5.90	1.40
Low income	1.95	1.99	1.02	1.12	1.98	1.77	1.00
Lower middle income	3.55	2.01	0.56	3.15	2.12	0.67	1.06
Middle income	2.73	1.67	0.61	2.70	1.72	0.64	1.03
Upper middle income	1.37	2.18	1.59	0.71	2.40	3.39	1.10
World	1.92	1.38	0.72	1.81	1.12	0.62	0.81

II. Estimating the welfare cost of volatility

A. Method

Most quantifications of the welfare cost of volatility are variants of the original Lucas exercise, consisting in estimating the expected welfare of an uncertain (risky) consumption path and comparing it with the welfare produced by a less risky (deterministic) consumption path. Since the (risk-averse) consumer would prefer the deterministic consumption, the risky one would have to be multiplied by a factor (greater than one), such that the welfare implicit in the choice of the risky path is equal to the welfare level of the deterministic consumption. Lucas estimated that to equalize the welfare in both cases, the compensating factor had to be $0.5\,\gamma\sigma^2$. This expression is the basis for computing the welfare cost of consumption volatility, which depends on the degree of risk aversion (γ) and the amount of risk (σ).

Lucas estimated this cost for the US economy at less than one twentieth of one percent of consumption. This extremely low value could be explained partly by the low risk aversion parameter and the low volatility of the US economy. His exercise was based on a simple constant relative risk aversion (CRRA) preference function, and on a stochastic consumption stream with deterministic mean but subject to random transitory disturbances. Later contributions to this literature focused on considering different types of utility functions as well as stochastic consumption process with permanent shocks(Obstfeld, 1994). The revised estimates were substantially larger than Lucas', but still below or around one percent of consumption.

Most quantifications of this cost refer to the US economy and there are some for OECD economies. There are very few done for developing countries, and we will use their results to benchmark the estimates presented in this note. The particular application considered here, by Athanasoulis and van Wincoop (AW) (2000), is based on comparing the welfare implicit in a country's risky consumption path with the welfare level implicit in the minimum risk consumption. The minimum risk consumption path is that which results when the country has diversified completely the country-specific risk, emanating from country-specific shocks. The minimum volatility is global, and is risk that cannot be diversified, This level is achieved when there is complete risk sharing across countries. The risk averse consumer would prefer the risk-sharing consumption path, so the individual country consumption stream would have to be multiplied by the compensating factor such that both welfare levels are equalized.

AW find that an approximation to this compensating factor is given by $0.5\,\gamma\Omega(\sigma_i^2 - \sigma_W^2)$. This formula is similar to Lucas', but now the amount of risk is the difference between the individual country volatility and the world volatility[23], that is multiplied by a series of weights (Ω) that depend on the difference between the risk-free rate of interest and the risk-adjusted global growth rate.[24]

B. Data-

Given the sizes of these variables, the key determinants of the cost will be the RRA parameter and the difference in volatility of endowment (consumption or output) in the country and the world aggregate. In general, public finance applications use a RRA ranging from 1 to 4 (Lucas, 2003). The two applications to developing countries used to benchmark our results use different values. De Ferranti, et. al (2000) use 3, while Pallage and Robe (2003) use a range of values from 1.5 to 10. Athanasoulis and van Wincoop use 3, but report results for the representative country rather than for individual countries of their sample. To facilitate comparisons with previous papers we use a RRA parameter of 3, noting that the estimate of the cost will be proportional to this value and facilitates sensitivity analysis to the particular assumed value..

The assumption of a common RRA parameter for all countries might be somewhat restrictive, but there is little else to do. Probably risk aversion could be decreasing with the level of GDP, just as risk aversion at the household level is found to be decreasing with the level of endowment in Italy (Guiso and Paiella, 2003). Other papers have found that, across households, RRA depends on the level of education, health and employment status of the head of household. (Eisenhauer and Ventura, 2003). Hence, it seems natural to assume different RRA across countries, but there is very little work in this area despite its importance determining saving and expenditure.

As a proxy for the volatility, we used the standard deviation (σ) of both GDP and consumption for each country during 1991-2005 (Table 3). Hence, the difference between the country volatility and global volatility is a constant, in contrast to the original AW paper where it is varying. This leads to a modification of the original sum of weighted factors to a simple sum of discount factors.

The discount factors are estimated for a 35-year horizon[25], and the risk free real interest rate is estimated as in AW: to a base value of 0.8 percent (from Mehra and Prescott), a term premium is added, constructed as the average difference between the *t* year bond yield and the one year bond yield, from 1961 to 2006, as reported in the Gurkaynak et.al (2006) data base.

C. The sample of countries

The sample was selected to facilitate comparison with previous estimates. The LAC countries are a subset of those reported in de Ferranti et.al., while the African countries are a subset of those in Pallage and Robe, plus 4 other countries selected to enlarge the African sample.

---

[23] AW estimate country and world output volatility from growth regressions, as the variance in the growth innovation. Estimates of the volatility are obtained running regressions at different horizons, hence the residual risk varies as the forecasting horizon changes.

[24] The weights are given by $\Omega_t = \sum_{t=1}^{T} \frac{e^{-(r_t - \mu_t)}}{\sum_{s=1}^{T} e^{-(r_s - \mu_s)}}$ where *r* is the risk free rate and μ is the risk-adjusted growth rate

[25] As in AW and de Feranti et.al.

Table 3
Output and Consumption per capita growth in a sample of countries, 1991-2005

	GDP			Consumption			
	$\bar{x}$	σ	$\sigma / \bar{x}$	$\bar{x}$	σ	$\sigma / \bar{x}$	σ con/ σ gdp
Algeria	1.02	2.90	2.85	-0.08	4.57	-57.30	1.58
Botswana	3.70	2.09	0.57	2.18	4.58	2.10	2.19
Egypt	2.31	1.37	0.59	2.07	1.40	0.68	1.02
Lesotho	2.44	3.19	1.31	-0.42	7.90	-18.85	2.48
Malawi	1.28	6.73	5.26	3.25	15.94	4.90	2.37
Mauritius	3.70	0.95	0.26	3.74	1.09	0.29	1.15
Morocco	1.42	5.24	3.69	0.66	6.07	9.21	1.16
Nigeria	1.12	2.73	2.44	4.38	24.85	5.67	9.12
South Africa	0.80	2.66	3.32	1.43	2.80	1.96	1.05
Sudan	3.60	1.62	0.45	1.72	9.94	5.79	6.12
Tunisia	3.24	1.85	0.57	3.09	1.63	0.53	0.88
Uganda	2.96	2.11	0.71	2.13	3.23	1.51	1.53
Zimbabwe	-2.54	6.01	-2.36	-2.85	12.47	-4.38	2.08
Argentina	2.71	6.72	2.48	2.52	8.29	3.29	1.23
Bolivia	1.34	1.41	1.05	0.88	1.08	1.23	0.77
Brazil	1.03	1.96	1.90	1.58	5.63	3.57	2.87
Chile	4.26	3.04	0.71	4.80	3.37	0.70	1.11
Colombia	1.04	2.47	2.37	0.75	3.60	4.83	1.46
Costa Rica	2.38	2.62	1.10	2.27	3.32	1.46	1.27
Ecuador	1.17	3.11	2.66	1.61	3.78	2.34	1.22
El Salvador	1.77	1.94	1.09	2.20	2.61	1.19	1.35
Guatemala	1.21	1.01	0.83	1.64	0.67	0.41	0.67
Mexico	1.51	3.18	2.11	1.76	4.00	2.28	1.26
Nicaragua	1.09	2.41	2.22	4.27	9.67	2.26	4.02
Paraguay	-0.58	1.84	-3.19	0.49	4.52	9.15	2.45
Peru	2.33	3.62	1.55	1.80	3.06	1.70	0.84
Uruguay	1.75	5.67	3.25	2.44	8.37	3.44	1.48
Venezuela	0.38	7.00	18.24	-0.09	6.33	-68.69	0.90

Two caveats when interpreting results:

1) the volatility of consumption is greater than for GDP, and in some cases there are wide discrepancies in the same country. This is particularly relevant in Africa. For instance, in Algeria, Malawi, Sudan and Nigeria, the volatility of consumption is more than twice that of GDP. In LAC, Brazil, Nicaragua and Paraguay show the same discrepancy.
2) Some countries have relatively similar volatility (σ) but different average growth rates. Such is the case of GDP growth in Chile, Mexico and Ecuador, where σ is around 3, but Chile's average rate is two and three times larger, respectively, than in the other two. To verify the pervasiveness of this phenomenon, we plotted the volatility measures and the averages for consumption and GDP growth for the 28 countries in the sample (Graph 1), rejecting the existence of any statistically significant relationship between volatility and averages in consumption and moderately significant, but negative, in GDP growth.

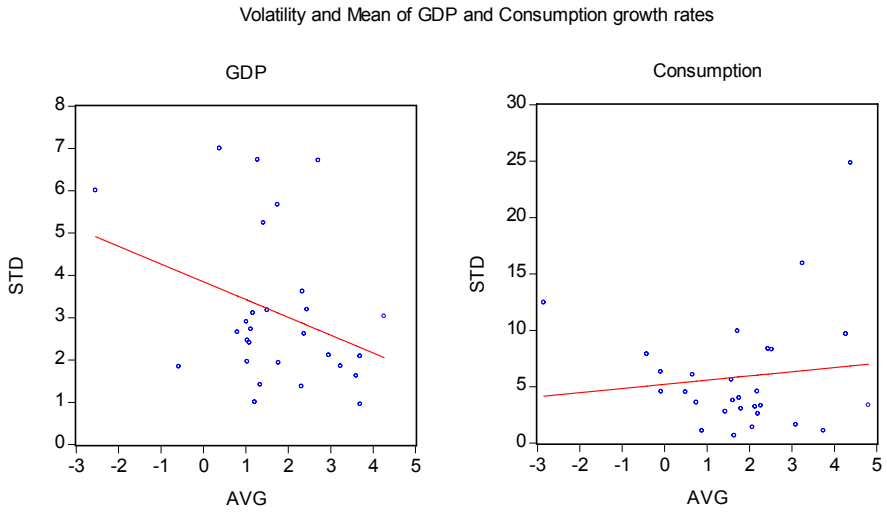


Table 4
Costs of Output and Consumption Volatility*

	Output	Consumption
Algeria	0.033	0.090
Botswana	0.015	0.090
Egypt	0.004	0.007
Lesotho	0.041	0.272
Malawi	0.195	1.117
Mauritius	0.000	0.004
Morocco	0.116	0.160
Nigeria	0.029	2.704
South Africa	0.027	0.033
Sudan	0.008	0.431
Tunisia	0.011	0.010
Uganda	0.016	0.044
Zimbawe	0.154	0.680
Argentina	0.194	0.244
Bolivia	0.005	0.004
Brazil	0.013	0.137
Chile	0.036	0.048
Colombia	0.023	0.055
Costa Rica	0.026	0.047
Ecuador	0.039	0.061
El Salvador	0.013	0.028
Guatemala	0.001	0.000
Mexico	0.040	0.069
Nicaragua	0.022	0.408
Paraguay	0.011	0.088
Peru	0.054	0.039
Uruguay	0.137	0.305
Venezuela	0.211	0.174

*Results indicate the permanent percentage increase in output or consumption that would produce a welfare gain equivalent to the reducing volatility to the global level

To verify the accuracy of the estimates based on this simplified formula, we compare our results with previous paper's estimates. In addition to the developing countries listed in Table 4, we also estimated the cost of volatility for a group of OECD countries given that they are the focus of most of the literature.

Table 5 compares the estimates based on the simplified formula with Pallage and Robe's (2003) for a sample of African countries and the US. The sample period was modified to match PR's (1968-1996). The authors provide a range of estimates that vary widely depending on the specific values of the relative risk aversion (RRA) parameter and the assumed value of the inter-temporal elasticity of substitution. We consider values of RRA of the order assumed in this note between 2.5 and 5, and lower values of the elasticity of inter-temporal substitution to report the lower and upper bounds of the PR estimates. The median of the cost of consumption volatility for the whole sample of African countries in the PR paper oscillates between 4 and 24 percent.

Table 5
Comparison with Pallage and Robe (PR) estimates of cost of consumption volatility 1968-1996

	Our estimate	PR estimates	
		lower	upper
Algeria	0.174	0.004	0.16
Egypt	0.023	0.011	0.043
Lesotho	0.156	0.212	0.404
Malawi	0.491	0.212	0.697
Mauritius	0.189	0.041	0.241
Morocco	0.099	0.017	0.045
South Africa	0.019	0.024	0.089
Sudan	0.101	0.163	0.407
Tunisia	0.074	0.012	0.053
USA	0.006	0.004	0.015

Table 6 compares the simplified estimates of consumption volatility for a group of LAC countries with De Ferranti et. al. (2000). The median cost of consumption volatility for the sample of LAC countries in the de Ferranti exercise is 7 percent.

Table 6 Comparison of Welfare cost of volatility in LAC countries, 1991-1999			
	our estimates		de Ferranti et. al.
	Output	Consumption	
Argentina	0.139	0.127	0.096
Bolivia	0.008	0.003	0.000
Brazil	0.038	0.177	0.070
Chile	0.052	0.067	0.037
Colombia	0.036	0.081	0.010
Costa Rica	0.031	0.043	0.030
Ecuador	0.038	0.042	0.000
El Salvador	0.012	0.241	0.137
Guatemala	0.000	0.003	0.003
Mexico	0.054	0.094	0.085
Nicaragua	0.037	0.890	0.634
Paraguay	0.007	0.140	0.460
Peru	0.113	0.117	0.065
Uruguay	0.054	0.147	0.086
Venezuela	0.093	0.078	0.065

Table 7 compares the simplified estimates of consumption volatility costs in a group of OECD counties with the van Wincoop (1994) estimates for the period 1970-1988. The median of the van Wincoop estimates oscillates between 1.1 percent and 2.7[26]. Table 8 updates the cost estimate with a more recent sample and and expanded sample.

Table 7 Comparison of consumption volatility cost in a sample of OECD countries, 1970-1988			
	our	van Wincoop	
		lower	upper
Australia	0.001	0.008	0.019
Canada	0.016	0.012	0.030
Denmark	0.031	0.013	0.032
Finland	0.022	0.014	0.033
France	0.001	0.002	0.005
Iceland	0.265	0.067	0.200
Ireland	0.060	0.020	0.048
Japan	0.024	0.007	0.017
Norway	0.043	0.012	0.029
Spain	0.034	0.011	0.027
Switzerland	0.006	0.005	0.011
United Kingdom	0.031	0.006	0.014
United States	0.013	0.004	0.009

Conclusions

The simplified formula used in the present exercise provides a useful back-of-the-envelope approximation to estimate the cost of volatility. Despite its simplicity, the method yields results comparable to previous papers that use richer and more complicated models. Our estimates show that the developing countries in

[26] van Wincoop's range of values are those estimated with a RRA parameter of 4.

this sample could, by reducing volatility, increase the welfare of their citizens in magnitudes equivalent to a permanent increase in consumption of around 8 percent per year. For the African countries in our sample, the potential median gain is of the order of 9 percent, while for the LAC countries it is of the order of 6 percent. The issue of determinants of volatility has been explored elsewhere (De Ferrranti, 2000. Ftas et.al.), but most papers show that policy instability is a key factor.

A fruitful line of research would be to estimate RRA parameters for each country, based on household surveys. Gourinchas and Parker(2002) propose a useful methodology

Table 8
Welfare cost of output and consumption volatility
1991-2005

Australia	0.001	0.003
Canada	0.012	0.009
Denmark	0.004	0.012
Finland	0.042	0.037
France	0.003	0.004
Germany	0.004	0.005
Iceland	0.031	0.080
Ireland	0.030	0.019
Italy	0.001	0.010
Japan	0.004	0.003
Norway	0.005	0.004
Spain	0.003	0.007
Switzerland	0.004	0.003
United Kingdom	0.004	0.008
United States	0.004	0.005
China	0.011	0.018
India	0.015	0.016
Indonesia	0.095	0.072
Korea	0.054	0.140
Russia	0.230	0.138
Saudia Arabia	0.028	0.027
Turkey	0.106	0.115

.

References for Appendix 1

Agenor, P. J. Mc Dermott, and E. Prasad (2000) "Macroeconomic fluctuations in developing countries: some stylized facts" *The World Bank Economic Review* 14, 2

Athanasoulis, S. and E. van Wincoop (2000) "Growth uncertainty and risksharing", *Journal of Monetary Economics*,45 pp. 477-505

De Ferranti, D., G. Perry, I. Gill and L. Serven (2000) *Securing our Future in a Global Economy*. The World Bank.

Eisenhauer, J. and L. Ventura "Survey measures of risk aversion and prudence", *Applied Economics*, 35, 1477-1484

Gourinchas, P. and J. Parker (2002) "Consumption over the life cycle", *Econometrica*, vol 70, No. 1

Guiso and Paiella (2003) "Risk aversion, wealth and background risk". Mimeo. Bank of Italy

Gurkaynak, R., B. Sack, and J. Wright (2006) "The US Treasury yield curve: 1961 to the present", Finance and Economics Discussion Series, Federal Reserve Board 2006-28. June.

Lucas, R. (1987) *Models of Business Cycles*

Lucas, R. (2003) "Macroeconomic priorities", *The American Economic Review*, March 2003

Neumayer, P. and F. Perri (2005) "Business cycles in emerging economies: the role of interest rates" *Journal of Monetary Economics* 52 pp. 345-380

Pallage, S. and M. Robe (2003) On the welfare cost of economic fluctuations in developing countries, *International Economic Review* Vol. 44, No. 2, pp. 677-998

Van Wincoop, E. (1994) "Welfare gains from international risksharing" *Journal of Monetary Economics*, 34 pp. 175-200

Appendix 2
Simple analytics of the growth effect of additional productive spending finaced with taxation (Adam-Bevan, 2005)

This is a very brief summary of the Adam Bevan model that captures the effect of additional taxes to fund higher government **productive spending**. The main point is to highlight that the effect of additional spending on growth will depend on the ratio of productivities of the public sectors and private sectors, the country's debt level, and other parameters such as the average tax rate and the discount rate. The model is quite rich and the reader is urged to consult the original paper.

The preferences:
It is a two-period model with a log-utility function described as:

$U= b\ln c_1 + (1-b) \ln c_2$

Technology:

The representative firm's production function incorporates the possibility that marginal (α) and average (β) products of private capital differ, and includes positive externalities of the overall capital intensity and productive public expenditure on the representative firm's output (y_i)

$$y_i = A^{\alpha+\beta} l_i^{1-\alpha} k_i^{\alpha} (K/L)^{\beta} (G_p/L)^{1-\alpha-\beta}$$

For the aggregate economy, output is given by:

$$Y= A^{\alpha+\beta} K^{\alpha+\beta} G_p^{1-\alpha-\beta} = AK \gamma_p^{(1-\alpha-\beta)/(\alpha+\beta)}$$

Where $\gamma_p = G_p/Y$

Government Budget:

Abstracting from seignoriage, external grants and external debt (all considered in the original paper but complicate matters unnecessarily for present purposes), the government's budget (all variables as a share of GDP) can be expressed as:

$$\tau_t = \gamma_{pt} + \gamma_{ut} + \left(\frac{(1+r_t)}{(1+g)_t} \Delta_t \right) - \Delta_{t+1}$$

Where $\gamma_p = G_p/Y$= *productive expenditure*
$\gamma_u = Gu/Y$ = *unproductive expenditure*

Δ_t = debt stock at end of period t
r= rate of interest, g= growth rate of output

The growth rate of output between geriods t and t+1, g_{t+1} is given by:

$$g_{t+1} = Y_{t+1}/Y_t - 1 = [AK_{t+1}\, \gamma_{pt+1}^{(1-\alpha-\beta)/(\alpha+\beta)} / Y_t] - 1$$

$$= A\,[(1-b)\,(1-\alpha)\,(1-\tau_t)-\Delta_{t+1}]\, \gamma_{pt+1}^{(1-\alpha-\beta)/(\alpha+\beta)} - 1$$

From the above expression, Adam_Bevan show that a tax rate increase to finance productive spending will be growth enhancing if :

$$(1-b)\,(1-\alpha)\,[(1-\tau_t) - \gamma_p(\alpha+\beta)/(1-\alpha-\beta)] \geq \Delta_{t+1}$$

With the above expression form the Adam-Bevan paper, we can gauge the order of magnitude of the impact of additional productive expenditure, financed with taxation, on growth. Hence we need parameter values for b (preference parameter), the marginal productivity of private capital (α), the ratio of productive expenditure to GDP (γ_p), the average elasticity of output with respect to private capital $(\alpha+\beta)$, the elasticity of output with respect to productive expenditures $(1-\alpha-\beta)$, and the level of public debt to GDP (Δ_{t+1})

Starting with the elasticity of output with respect to productive expenditure $(1-\alpha-\beta)$, discussion in the main text indicates that it is between .1 and .3. For the purpose of this exercise we will consider the upper range of values of around .4.

The elasticity of output with respect to private capital is obtained as a product of the return to capital times the private capital to output ratio. Assuming a real return of 20%, and a capital to output ratio of 2, then the elasticity , β, is .4. With $\beta=.4$ and $(1-\alpha-\beta)=.4$, then $\alpha=.2$.

The average tax rates of the economy can fluctuate from .2 in developing countries, to .4 in developed nations.

The debt to GDP ratio can be around 40%, while in developed nations might be around 80%. The ratio of productive spending to GDP is assumed around 15%.

Finally, the preference parameter, b, can be derived from observed savings rates. The functional form implies a consumption of bw and a savings level of (1-b)w, with a wage rate of w[27]. With savings rates of about 20%, b would be around .8.

With there parameter values the response of a change in a percentage point of productive expenditure can be estimated at -.3. The negative response is fairly robust to most parameter variations. When the debt ratio increases to 60% of GDP, the negative impact is larger, increasing to -.7.

[27] This is from the footnote 8 in the Adam-Bevan paper.

Appendix 3
Methods for Measuring Technical Efficiency (following Herrera and Pang 2006)

The origin of the modern discussion of efficiency measurement dates back to Debreu (1951) and Farell (1957), who identified two different ways in which productive agents could be inefficient: one, they could use more inputs than technically required to obtain a given level of output, or two, they could use a sub-optimal input combination given the input prices and their marginal productivities. The first type of inefficiency is termed technical inefficiency while the second one is known as allocative inefficiency.

These two types of inefficiency can be represented graphically by means of the unit isoquant curve in Figure1. The set of minimum inputs required for a unit of output lies on the isoquant curve YY'. An agent's input-output combination defined by bundle P produces one unit of output using input quantities X_1 and X_2. Since the same output can be achieved by consuming less of both inputs along the radial back to bundle R, the segment RP represents the inefficiency in resource utilization. The technical efficiency (TE), input-oriented, is therefore defined as TE = OR/OP. Furthermore, the producer could achieve additional cost reduction by choosing a different input combination. The least cost combination of inputs that produces one unit of output is given by point T, where the marginal rate of technical substitution is equal to the input price ratio. To achieve this cost level implicit in the optimal combination of inputs, input use needs to be contracted to bundle S. The input allocative efficiency (AE) is defined as AE = OS/OR.

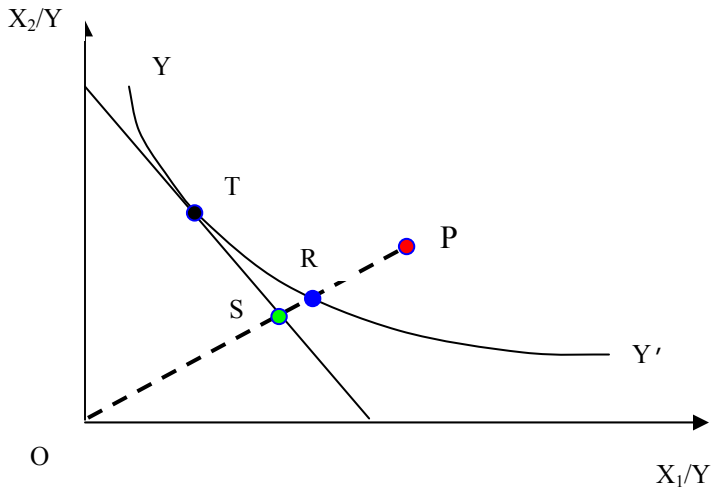


Figure 1 Technical and Allocative Inefficiency

The focus of this paper is measuring technical efficiency, given the lack of comparable input prices across the countries. This concept of efficiency is narrower than the one implicit in social welfare analysis. That is, countries may be producing the wrong output very efficiently (at low cost). We abstract from this consideration (discussed by Tanzi 2004), focusing on the narrow concept of efficiency.

Numerous techniques have been developed over the past decades to tackle the empirical problem of estimating the unknown and unobservable efficient frontier (in this case the isoquant YY").

These may be classified using several taxonomies. The two most widely used catalog methods into parametric or non-parametric, and into stochastic or deterministic. The parametric approach assumes a specific functional form for the relationship between the inputs and the outputs as well as for the inefficiency term incorporated in the deviation of the observed values from the frontier. The non-parametric approach calculates the frontier directly from the data without imposing specific functional restrictions. The first approach is based on econometric methods, while the second one uses mathematical programming techniques. The deterministic approach considers all deviations from the frontier explained by inefficiency, while the stochastic focus considers those deviations a combination of inefficiency and random shocks outside the control of the decision maker.

This paper uses non-parametric methods to avoid assuming specific functional forms for the relationship between inputs and outputs or for the inefficiency terms. A companion paper will explore the parametric approach, along the lines proposed by Greene (2003). The remainder of the section briefly describes the two methods: the Free Disposable Hull (FDH) and the Data Envelopment Analysis (DEA)

The FDH method imposes the least amount of restrictions on the data, as it only assumes free-disposability of resources. Figure 2 illustrates the single-input single-output case of FDH production possibility frontier. Countries A and B use input X_A and X_B to produce outputs Y_A and Y_B, respectively. The *input efficiency* score for country B is defined as the quotient X_A/X_B. The *output efficiency* score is given by the quotient Y_B/Y_A. A score of one implies that the country is on the frontier. An input efficiency score of 0.75 indicates that this particular country uses inputs in excess of the most efficient producer to achieve the same output level. An output efficiency score of 0.75 indicates that the inefficient producer attains 75 percent of the output obtained by the most efficient producer with the same input intake. Multiple input and output efficiency tests can be defined in an analogous way.

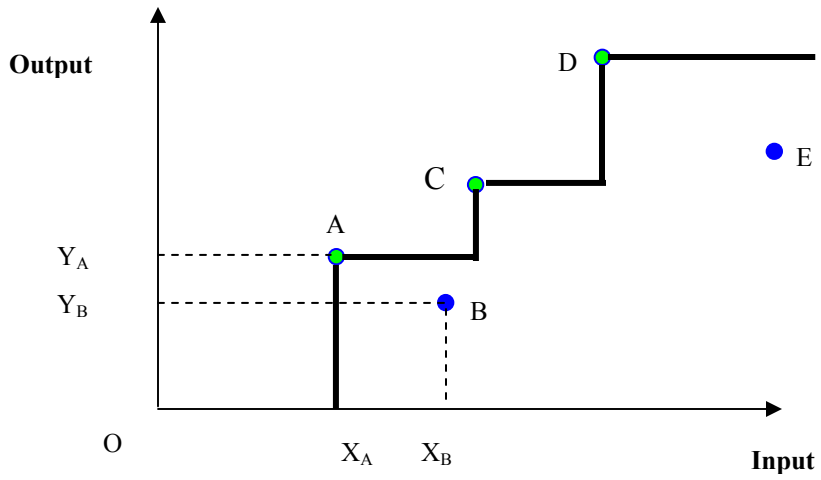


Figure 2 Free Disposal Hull (FDH) production possibility frontier

The second approach, Data Envelopment Analysis (DEA), assumes that linear combinations of the observed input-output bundles are feasible. Hence it assumes convexity of the production set to construct an envelope around the observed combinations. Figure 3 illustrates the single input-single output DEA production possibility frontier. In contrast to the vertical step-ups of FDH frontier, DEA frontier is a piecewise linear locus connecting all the efficient decision-making units (DMU). The feasibility assumption, displayed by the piecewise linearity, implies that the efficiency of C, for instance, is not only ranked against the real performers A and D, called the peers of C in the literature, but also evaluated with a virtual decision maker, V, which employs a weighted collection of A and D inputs to yield a virtual output. DMU C, which would have been considered to be efficient by FDH, is now lying below the variable returns to scale (VRS, further defined below) efficiency frontier, XADF, by DEA ranking. This example shows that FDH tends to assign efficiency to more DMUs than DEA does. The input-oriented technical efficiency of C is now defined by TE = YV/YC.

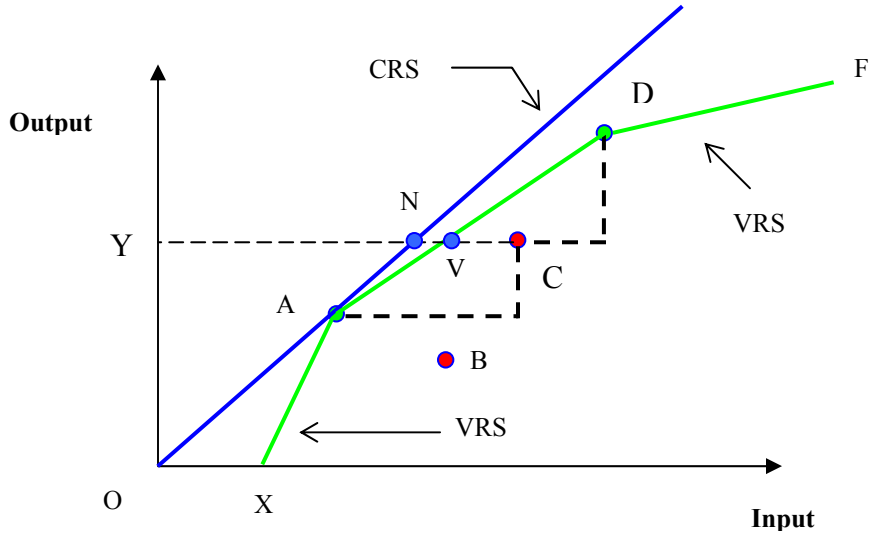


Figure 3 DEA production possibility frontier

If constant returns to scale (CRS) characterize the production set, the frontier may be represented by a ray extending from the origin through the efficient DMU (ray OA). By this standard, only A would be rated efficient. The important feature of the XADF frontier is that this frontier reflects variable returns to scale. The segment XA reflects locally increasing returns to scale (IRS), that is, an increase in the inputs results in a greater than proportionate increase in output. Segments AD and DF reflect decreasing returns to scale. It is worth noticing that constant returns to scale technical efficiency (CRSTE) is equal to the product of variable returns to scale technical efficiency (VRSTE) and scale efficiency (SE). Accordingly, DMU D is technically efficient but scale inefficient, while DMU C is neither technically efficient nor scale efficient. The scale efficiency of C is calculated as YN/YV. For more detailed exploration of returns to scale, readers are referred to Charnes, Cooper, and Rhodes (1978) and Banker, Charnes, and Cooper (1984), among others.

The limitations of the non-parametric method derive mostly from the sensitivity of the results to sampling variability, to the quality of the data and to the presence of outliers. This has led recent literature to explore the relationship between statistical analysis and non-parametric methods

(Simar and Wilson, 2000). Some solutions have been advanced. For instance, confidence intervals for the efficiency scores can be estimated using asymptotic theory in the single input case (for input-efficiency estimators) or single-output (in the output efficiency) case, given these are shown to be maximum likelihood estimators (Banker, 1993 and Goskpoff, 1996). For multiple input-output cases the distribution of the efficiency estimators is unknown or quite complicated and analysts recommend constructing the empirical distribution of the scores by means of bootstrapping methods (Simar and Wilson, 2000). Other solutions to the outlier or noisy data consist in constructing a frontier that does not envelop all the data point, building an expected minimum input function or expected maximum output functions (Cazals, Florens and Simar, 2002, and Wheelock and Wilson, 2003). Another limitation of the method, at least in the context in which we will apply it, is the inadequate treatment of dynamics, given the lag between input consumption (public expenditure) and output production (health and education outcomes).

Appendix 4
The rate of return of public projects and the marginal cost of public funds (based on Bevan, 2007)

After analyzing the underlying rationale that justifies public intervention through specific projects or programs, the available options have to be scrutinized and the costs and benefits have to be evaluated. Government expenditure may produce returns of different types: they may be direct and financial, indirect and financial, or indirect and social. The benefit-cost analysis should contemplate these differences in the impact the project will have on the future. Whether the project is "productive" and pays for itself fully in the future with the additional revenue it generates, or if it is "unproductive" and will not generate any future public revenue (all the return is private) will affect the estimation of the required rate of return to ensure that the net social benefit will be positive.

To illustrate the differences between types of projects, consider three variations of an infinitely lived project. The cases differ according to whom the benefit accrues. In all cases, the project requires an initial cash investment of 1, which is financed by borrowing; project *i* then yields a constant flow of benefit, b_i and has a constant cash operating cost, c_i. The discount rate, r, is constant and common to both government and the representative private agent. The government receives revenue from an income tax at rate τ. This inflicts a deadweight loss on the private agent at an average rate θ_A and a higher marginal rate θ_M per unit of revenue raised.

Case 1: Direct financial benefits fully appropriated by the government covering the full cost of the project

This has a benefit flow which is all appropriable by government, for example via user charges. To the extent that this appropriability property also held for a private investor, there is no necessity for the public sector to undertake the project, which could be left to the private sector, but it provides a useful benchmark. The present value criterion is to undertake the project if:
$\sum_{1}^{\infty} \frac{(b_1 - c_1)}{(1+r)^t} \geq 1$. Summing the left hand side yields $\frac{(b_1 - c_1)}{r}$, so the benefit-cost test in this case requires that:

$$b_1 \geq r + c_1 \qquad (1)$$

Case 2: Indirect Financial Benefits

This project produces a benefit flow that accrues entirely as additional private income. However, due to taxation, public revenue increases and private benefit is reduced to $b_2\left[1-\tau(1+\theta_A\right]$. If the additional public revenue is not enough to fully pay for the project cost ($b_2\tau < r + c_2$), then, additional taxation equal to $r + c_2 - b_2\tau$ must be

levied. This lowers private benefit further by $(r + c_2 - b_2\tau)(1+\theta_M)$. Hence social net benefit is given by:
$b_2\left[1-\tau(1+\theta_A\right]-(r+c_2-b_2\tau)(1+\theta_M)=b_2\left[1+\tau(\theta_M-\theta_A)\right]-(r+c_2)(1+\theta_M)$.
For the project to be worth undertaking, this must be non-negative. Hence the benefit-cost test in this case is:

$$b_2 \geq \frac{(r+c_2)(1+\theta_M)}{1+\tau(\theta_M-\theta_A)} \qquad (2)$$

Case 3: Indirect Social Benefits

This has a benefit flow which affects purely private utility, without any productive impact so that none of the benefit is monetary and hence the government does not capture any of the project's return. Hence, all expenditure must be fully funded by additional taxation. The additional revenue flow must equal $(r+c_3)$, imposing a cost on the private agent of $(r+c_3)(1+\theta_M)$. The private (and social) net benefit flow is $b_3-(r+c_3)(1+\theta_M)$. For this to be positive, the non-negativity benefit-cost condition becomes:

$$b_3 \geq (r+c_3)(1+\theta_M) \qquad (3)$$

It is apparent that, for given c, $b_3 > b_2 > b_1$.

As an illustration of the benchmarks that these simplifications produce, Bevan (2007) considers a case in which $r = 0.05$, and $c = 0.05$, $\tau = 0.20, \theta_A = 0.20, \theta_M = 0.40$ and obtains $b_1 \geq 0.10, b_2 \geq 0.1346, b_3 \geq 0.14$, judged to be a reasonable range.